

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:27:12 : Search time 35 Seconds  
(without alignments)  
34.264 Million cell updates/sec

Title: US-09-641-528A-31040  
Perfect score: 44  
Sequence: 1 ATLERTVEY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	158	22	Human papillomavir
2	32	72.7	159	22	Staphylococcus aur
3	32	72.7	162	22	Staphylococcus aur
4	32	72.7	383	21	Fatty acid modifi
5	32	72.7	383	21	Impatiens balsamin
6	32	72.7	438	23	Streptococcus poly
7	32	72.7	663	23	Streptococcus poly
8	32	72.7	663	23	Streptococcus poly
9	32	72.7	692	23	Streptococcus poly
10	32	72.7	1242	21	Human colon cancer

11	32	72.7	1428	22	ABB65364	Drosophila melanog
12	31	70.5	99	21	AAG25796	Arabidopsis thalia
13	31	70.5	103	21	AAG25795	Arabidopsis thalia
14	31	70.5	113	23	ABB89417	Human polypeptide
15	31	70.5	128	21	AAG04807	Arabidopsis thalia
16	31	70.5	132	21	AAG04806	Arabidopsis thalia
17	31	70.5	135	22	ABB69313	Drosophila melanog
18	31	70.5	157	21	AAG49165	Arabidopsis thalia
19	31	70.5	161	21	AAG25794	Arabidopsis thalia
20	31	70.5	161	21	AAG49164	Arabidopsis thalia
21	31	70.5	181	22	ABB60217	Burkholderia cepac
22	31	70.5	190	21	AAG04805	Arabidopsis thalia
23	31	70.5	202	21	AAG49147	Arabidopsis thalia
24	31	70.5	206	21	AAG49146	Arabidopsis thalia
25	31	70.5	219	21	AAG49163	Arabidopsis thalia
26	31	70.5	229	10	AAP93524	Complete sequence
27	31	70.5	229	21	AAY67943	Human type IV coll
28	31	70.5	229	23	AAU75587	Human type IV coll
29	31	70.5	260	20	AAU31991	Type IV collagen N
30	31	70.5	260	21	AAY97553	Human alpha(IV)NC
31	31	70.5	264	21	AAG49145	Arabidopsis thalia
32	31	70.5	353	22	AAG81770	S. epidermidis ope
33	31	70.5	406	21	AAB58169	Lung cancer associ
34	31	70.5	448	22	AAG82317	S. epidermidis ope
35	31	70.5	448	23	ABP38210	Staphylococcus epi
36	31	70.5	457	17	AAR86803	S. clavuligerus ly
37	31	70.5	683	23	ABP40731	Staphylococcus epi
38	31	70.5	728	22	ABG16004	Novel human diagno
39	31	70.5	1452	22	AAU78827	Human protein SEQ
40	31	70.5	1621	22	AAU79811	Human protein SEQ
41	31	70.5	1669	22	AAU40863	Human polypeptide
42	31	70.5	1669	23	ABP90760	Human Tumour Endot
43	31	70.5	1669	23	ABP57334	Mouse ischaemic co
44	31	70.5	1672	22	AAU39077	Human polypeptide
45	30	68.2	70	23	ABP32150	Human ORF123 prot

## ALIGNMENTS

RESULT 1  
AAB98440  
ID AAB98440 standard; Protein; 158 AA.

XX

AC AAB98440;

XX

DT 22-AUG-2001 (first entry)

XX

DE Human papillomavirus protein HPV45 E6.

XX

DE Human papillomavirus; human leukocyte antigen; HLA; immune response;

KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;

KW neoplastic growth; antiviral.

XX

OS Human papillomavirus.

XX

PN WO200141799-A1.

XX

PD 14-JUN-2001.

XX

PF 11-DEC-2000; 2000WO-US33549.

XX

PR 10-DEC-1999; 99US-0172705.

PR 15-AUG-2000; 2000US-0641528.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Sette A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HM;

XX

DR WPI; 2001-381497/40.

XX

PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for treating HPV infections -

XX Disclosure; Page 25; 756pp; English.

PS The present invention describes an isolated prepared human papillomavirus

CC (HPV) epitope (I). (I) has antiviral activity, and can be used in

CC vaccine production. Peptides and corresponding nucleic acid compositions

CC from the present invention are useful for stimulating an immune response

CC to HPV by stimulating the production of CTL or HTL responses,

CC specifically in the treatment or prophylaxis of HPV infection, in persons

CC who have not manifested symptoms e.g. genital warts or neoplastic growth.

CC The peptides can also be used in a tetramer staining assay to assess

CC peripheral blood mononuclear cells for the presence of antigen-specific

CC CTLs following exposure to a pathogen or immunogen, and as reagents to

CC evaluate immune recall responses or evaluate the efficacy of a vaccine.

CC The vaccine compositions are useful for removing warts or treating HPV

CC infections. The epitopes for inclusion in an epitope-base vaccine may

CC be selected from conserved regions of viral or tumour-associated

CC antigens, which reduces the likelihood of escape mutants, also

CC immunosuppressive epitopes that may be present in whole antigens can be

CC avoided with the use of epitope-base vaccines. An additional advantage

CC is the ability to combine selected epitopes (CTL and HTL) and to modify

CC the composition of the epitopes achieving enhanced immunogenicity, the

CC major benefit of the vaccine is that is safe and efficacious. AAB98391

CC to AAB98477 represent polypeptide sequences used in the exemplification

CC of the present invention.

XX

SQ Sequence 158 AA;

Query Match 100.0%; Score 44; DB 22; Length 158;

Best Local Similarity 100.0%; Pred. No. 0.32;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9

|||||||

Db 37 ATLERTEVY 45

RESULT 2

AAU33958

ID AAU33958 standard; Protein; 159 AA.

XX

AC AAU33958;

XX

DT 14-FEB-2002 (first entry)

XX

DE Staphylococcus aureus cellular proliferation protein #234.

XX

KW Antisense: prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US09180.

XX

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX

XX WPI; 2001-611495/70.

DR N-PSDB; AAS51817.

DR

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids .

XX

PS Example 3; Seq ID No 5454; 51lpp; English.

XX

CC The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 159 AA;

Query Match 72.7%; Score 32; DB 22; Length 159;

Best Local Similarity 75.0%; Pred. No. 89;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEV 8

|||||

Db 95 ATMERTEV 102

RESULT 3

AAU37090

ID AAU37090 standard; Protein; 162 AA.

XX

AC AAU37090;

XX

DT 14-FEB-2002 (first entry)

XX

DE Staphylococcus aureus cellular proliferation protein #1260.

XX

KW Antisense: prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US09180.

XX

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX

XX WPI; 2001-611495/70.

DR N-PSDB; AAS54949.

DR

XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 XX Example 3; Seq ID No 12683; 511pp; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 162 AA;

Query Match 72.7%; Score 32; DB 22; Length 162;  
 Best Local Similarity 75.0%; Pred. No. 91;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEV 8  
 II:III  
 Db 98 ATMEKTEV 105

#### RESULT 4

AAAY70401  
 ID AAY70401 standard; Protein; 383 AA.

XX AC AAY70401;

XX DT 21-JUN-2000 (first entry)

XX Fatty acid modifying enzyme encoded by I. balsamina ImpH8fad2 clone.

XX DE Clone ImpH8fad2; fatty acid modifying enzyme; transgenic plant;  
 XX KW seed oil; plant fatty acid; animal feed; carcass quality;  
 XX KW omega-6 fatty acid desaturase.

XX OS Impatiens balsamina.

XX PN WO200011176-A2.

XX PD 02-MAR-2000.

XX PF 16-AUG-1999; 99WO-US18562.

XX PR 20-AUG-1998; 98US-0097186.

XX PR 08-JUL-1999; 99US-0142756.

XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.

XX PI Cahoon EB, Carlson TJ, Hitz WD, Ripp KG;

XX XX WPI; 2000-246565/21.

XX DR N-PSDB; AA251591.

XX DNA sequence encoding plant fatty acid modifying enzyme useful for  
 PT producing seed oil containing altered levels of fatty acid with  
 PT conjugated double bond formation -

XX

PS Claim 2; Page 65-66; 87pp; English.

XX The present sequence is the instant fatty acid modifying enzyme  
 CC associated with conjugated double bond formation from seeds of  
 CC Impatiens balsamina encoded by the cDNA insert in clone ImpH8fad2. This  
 CC is structurally related to the omega-6 fatty acid desaturases. Chimeric  
 CC genes incorporating the corresponding nucleic acid fragments can be used  
 CC to create transgenic plants having altered lipid profiles. They are also  
 CC useful for producing seed oil-containing fatty acids. Fatty acids with  
 CC conjugated double bonds display high rates of oxidation. The animal feed  
 CC comprising the seeds and oil obtained from transformed host cells is  
 CC useful for improving carcass quality of an animal by supplementing it  
 CC within the animal diet.

XX SQ Sequence 383 AA;

Query Match 72.7%; Score 32; DB 21; Length 383;  
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLERTEV 9  
 I::III  
 Db 149 ASIERDEVY 157

#### RESULT 5

AAAY72819  
 ID AAY72819 standard; Protein; 383 AA.

XX AC AAY72819;

XX DT 31-MAY-2001 (first entry)

XX DE Impatiens balsamina ImpFad2H8 protein.

XX KW ImpFad2H8; fatty acid modifying enzyme; conjugated double bond; delta-9;  
 XX KW delta-12; transgenic plant; seed oil.

XX OS Impatiens balsamina.

XX PN WO200112800-A2.

XX PD 22-FEB-2001.

XX PF 15-AUG-2000; 2000WO-US22371.

XX PR 16-AUG-1999; 99US-0149050.

XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.

XX PI Cahoon EB, Hitz WD, Ripp KG;

XX DR WPI; 2001-218351/22.

XX Novel nucleic acid fragment encoding plant fatty acid modifying enzyme  
 PT associated with modification of delta-9 position of the fatty acid,  
 PT useful for creating transgenic plants having altered lipid profiles -

PS Example 5; Fig 2; 96pp; English.

XX The present invention relates to nucleic acid fragments encoding plant  
 CC fatty acid modifying enzymes associated with conjugated double bond  
 CC formation, comprising a modification of delta-9 position or a delta-12  
 CC position of fatty acids. The nucleic acid fragments are useful for  
 CC creating chimeric gene which is useful for creating transgenic plants  
 CC having altered lipid profiles. They are also useful for producing seed  
 CC oil containing fatty acids comprising a modification at delta-9 position  
 CC or delta-12 position. The seed is useful for improving the carcass  
 CC quality of an animal by supplementing a diet of the animal with animal  
 CC feed. The nucleic acid fragments are useful for isolating nucleic acid  
 CC fragments encoding homologous fatty acid modified enzymes from the same  
 CC or other plant species.  
 CC The present sequence is Impatiens balsamina ImpFad2H8 protein. It is

CC a fatty acid modifying enzyme associated with conjugated double bond  
 CC formation. It is also known to make eleostearic acid, a conjugated  
 CC fatty acid.

XX Sequence 383 AA;

SQ Query Match 72.7%; Score 32; DB 22; Length 383;  
 Best Local Similarity 66.7%; Pred. NO. 2.3e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLTERTEVY 9  
 |::|||

Db 149 ASIERDEVY 157

RESULT 6

ABP26430

ID ABP26430 standard; Protein; 438 AA.

XX AC ABP26430;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 2036.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 Tettelin H;

XX WPI: 2002-352536/38.

DR N-PSDB; ABN67061.

XX New Streptococcus protein for the treatment or prevention of infection  
 or disease caused by Streptococcus bacteria, such as meningitis, and  
 for detecting a compound that binds to the protein -

XX Claim 1; Page 3355; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 the specification. The proteins have antibacterial and antiinflammatory  
 activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 antibodies that bind (I) are used in the manufacture of medicaments for  
 the treatment or prevention of infection or disease caused by  
 Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 Nucleic acids encoding (I) is used to determine whether a compound binds to  
 biological sample. (I) is used to determine whether a compound binds to  
 Streptococcus that is prevented or treated may be meningitis. Nucleic  
 acid encoding (I) may be used to recombinantly produce (I) and may be  
 used in gene therapy. Antibodies to (I) are used for affinity  
 chromatography, immunoassays, and distinguishing/identifying  
 Streptococcus proteins.

XX SQ Sequence 438 AA;

Query Match 72.7%; Score 32; DB 23; Length 438;  
 Best Local Similarity 66.7%; Pred. NO. 2.7e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLTERTEVY 9  
 |::|||

Db 2 AKLERTDMY 10

RESULT 7

ABP27421

ID ABP27421 standard; Protein; 663 AA.

XX AC ABP27421;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 4018.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 Tettelin H;

XX WPI: 2002-352536/38.

DR N-PSDB; ABN68052.

XX New Streptococcus protein for the treatment or prevention of infection  
 or disease caused by Streptococcus bacteria, such as meningitis, and  
 for detecting a compound that binds to the protein -

XX Claim 1; Page 3557; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 the specification. The proteins have antibacterial and antiinflammatory  
 activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 antibodies that bind (I) are used in the manufacture of medicaments for  
 the treatment or prevention of infection or disease caused by  
 Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 Nucleic acids encoding (I) is used to determine whether a compound binds to  
 biological sample. (I) is used to determine whether a compound binds to  
 Streptococcus that is prevented or treated may be meningitis. Nucleic  
 acid encoding (I) may be used to recombinantly produce (I) and may be  
 used in gene therapy. Antibodies to (I) are used for affinity  
 chromatography, immunoassays, and distinguishing/identifying  
 Streptococcus proteins.

XX SQ Sequence 663 AA;

Query Match 72.7%; Score 32; DB 23; Length 663;  
Best Local Similarity 85.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
Db 486 TLERTEI 492

## RESULT 8

ABP27422

ID ABP27422 standard; Protein; 663 AA.

XX AC ABP27422;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 4020.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
XX KW antinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus pyogenes.

XX XN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Massignani V, Margarit Ros VI, Grandi G, Fraser C;  
XX PI Tettelin H;

XX DR WPI; 2002-352536/38.

XX DR N-PSDB; ABN68053.

XX XX New Streptococcus protein for the treatment or prevention of infection  
XX PT or disease caused by Streptococcus bacteria, such as meningitis, and  
XX PT for detecting a compound that binds to the protein -

XX PS Claim 1; Page 3557; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
XX CC the specification. The proteins have antibacterial and antinflammatory  
XX CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and  
XX CC antibodies that bind (I) are used in the manufacture of medicaments for  
XX CC the treatment or prevention of infection or disease caused by  
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
XX CC biological sample. (I) is used to determine whether a compound binds to  
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
XX CC used as a vaccine or diagnostic composition. The disease caused by  
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be  
XX CC used in gene therapy. Antibodies to (I) are used for affinity  
XX CC chromatography, immunoassays, and distinguishing/identifying  
XX CC Streptococcus proteins.

XX SQ Sequence 663 AA;

Query Match 72.7%; Score 32; DB 23; Length 663;  
Best Local Similarity 85.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
Db 486 TLERTEI 492

## RESULT 9

ABB53871

ID ABB53871 standard; Protein; 692 AA.

XX AC ABB53871;

XX DT 16-MAY-2002 (first entry)

XX DE Lactococcus lactis protein uvrB.

XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX OS Lactococcus lactis IL1403.

XX PN FR2807446-A1.

XX PD 12-OCT-2001.

XX PF 11-APR-2000; 2000FR-0004630.

XX PR 11-APR-2000; 2000FR-0004630.

XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX PI Boiotine A, Sorokine A, Renault P, Ehrlich SD;

XX DR WPI; 2002-043418/06.

XX PT New nucleotide sequence useful in the identification or Lactococcus  
XX PT lactis and related species -

XX PS Claim 6; SEQ ID No 573; 2504pp; French.

XX CC The present invention is related to a Lactococcus lactis nucleotide  
XX CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The  
XX CC nucleic acid sequence is useful in the detection and/or amplification of  
XX CC nucleic acid sequence, particularly to identify Lactococcus lactis or  
XX CC related species. The proteins of the invention are useful for the  
XX CC biosynthesis or biodegradation of a composition of interest. The  
XX CC invention helps research in lactic bacteria, particularly useful in the  
XX CC production of yogurt and cheese.

XX CC Note: The sequence data for this patent is based on equivalent patent  
XX CC WO200177334 (published 18-OCT-2001) which is available in electronic  
XX CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 692 AA;

Query Match 72.7%; Score 32; DB 23; Length 692;  
Best Local Similarity 85.7%; Pred. No. 4.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
Db 487 TLERTEI 493

## RESULT 10

AAB53371

ID AAB53371 standard; Protein; 1242 AA.

XX AC AAB53371;

XX DT 09-MAR-2001 (first entry)

XX DE Human colon cancer antigen protein sequence SEQ ID NO:911.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder.  
XX  
XX Homo sapiens.  
XX WO200005351-A1.  
XX  
XX PD 21-SEP-2000.  
XX  
XX PF 08-MAR-2000; 2000WO-US05883.  
XX  
XX PR 12-MAR-1999; 99US-0124270.  
XX  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Rosen CA, Ruben SM;  
XX  
XX DR WPI: 2000-587534/55.  
XX DR N-PSDB; AAC98128.  
XX  
XX PT Colon cancer associated gene sequences, referred to as colon cancer  
XX PT antigens, useful for the treatment, prevention, and diagnosis of colon  
XX PT disorders such as colon cancer -  
XX  
XX PS Claim 11: Page 1465-1470; 2104pp; English.  
XX  
XX CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
XX CC called human colon cancer antigens, given in AAB53234 to AAB54006. The  
XX CC human colon cancer antigens can have cytostatic, cardioactive, muscular;  
XX CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and  
XX CC vulnary, nephrotropic, antiinfective and antibacterial activities, and  
XX CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
XX CC proteins and antibodies to the proteins are useful for the prevention,  
XX CC treatment and diagnosis of colon disorders, such as colon cancer. The  
XX CC polynucleotides may be used in diagnostics, and research, such as for  
XX CC chromosome identification, and as hybridisation probes. The proteins  
XX CC may also be used to prevent diseases such as neural disorders, immune  
XX CC system disorders, muscular disorders, reproductive disorders,  
XX CC gastrointestinal disorders, wounds, renal disorders, infectious  
XX CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
XX CC AAB54007 represent sequences used in the exemplification of the present  
XX CC invention.  
XX  
XX SQ Sequence 1242 AA;  
Query Match 72.7%; Score 32; DB 21; Length 1242;  
Best Local Similarity 66.7%; Pred. No. 8.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATLETEVY 9  
II III::I  
Db 393 ATQERTDLY 401  
RESULT 11  
ABB5364  
ID ABB563364 standard; Protein; 1428 AA.  
XX  
XX AC ABB563364;  
XX  
XX DT 26-MAR-2002 (first entry)  
XX  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 22884.  
XX  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
XX OS Drosophila melanogaster.  
XX

PN WO200171042-A2.  
XX  
XX PD 27-SEP-2001.  
XX  
XX PF 23-MAR-2001; 2001WO-US09231.  
XX  
XX PR 23-MAR-2000; 2000US-191637P.  
XX PR 11-JUL-2000; 2000US-0614150.  
XX  
XX PA (PEKE ) PE CORP NY.  
XX  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX DR WPI: 2001-656860/75.  
XX DR N-PSDB; ABL09467.  
XX  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions -  
XX  
XX PS Disclosure; SEQ ID NO 22884; 21pp + Sequence Listing; English.  
XX  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX CC sequences (ABL01840-ABL16175) and the encoded proteins  
XX CC (ABB57737-ABB72072).  
XX CC The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX SQ Sequence 1428 AA;  
Query Match 72.7%; Score 32; DB 22; Length 1428;  
Best Local Similarity 87.5%; Pred. No. 9.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 TLETEVY 9  
III IIII  
Db 198 TLETEVY 205  
RESULT 12  
AAC25796  
ID AAG25796 standard; Protein; 99 AA.  
XX  
XX AC AAG25796;  
XX  
XX DT 17-OCT-2000 (first entry)  
XX  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30004.  
XX  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX OS Arabidopsis thaliana.  
XX  
XX PN EP1033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX PR 25-FEB-1999; 99US-0121825.  
XX PR 05-MAR-1999; 99US-0123180.  
XX PR 09-MAR-1999; 99US-0123548.  
XX PR 23-MAR-1999; 99US-0125788.  
XX PR 25-MAR-1999; 99US-0126264.  
XX PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139751.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.

PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999;	99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999;	99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999;	99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999;	99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	04-MAY-1999;	99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999;	99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	06-MAY-1999;	99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	06-MAY-1999;	99US-0132487.
PR	18-OCT-1999;	99US-0159584.	PR	07-MAY-1999;	99US-0132863.
PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999;	99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;	99US-0134218.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999;	99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999;	99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999;	99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160989.	PR	20-MAY-1999;	99US-0135124.
PR	25-OCT-1999;	99US-0161404.	PR	20-MAY-1999;	99US-0135353.
PR	25-OCT-1999;	99US-0161405.	PR	24-MAY-1999;	99US-0135629.
PR	25-OCT-1999;	99US-0161406.	PR	25-MAY-1999;	99US-0136021.
PR	26-OCT-1999;	99US-0161359.	PR	27-MAY-1999;	99US-0136392.
PR	26-OCT-1999;	99US-0161360.	PR	28-MAY-1999;	99US-0136782.
PR	26-OCT-1999;	99US-0161361.	PR	01-JUN-1999;	99US-0137222.
PR	28-OCT-1999;	99US-0161920.	PR	03-JUN-1999;	99US-0137528.
PR	28-OCT-1999;	99US-0161932.	PR	04-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161933.	PR	07-JUN-1999;	99US-0137724.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUN-1999;	99US-0138847.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUN-1999;	99US-0139119.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139452.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139453.
PR	29-OCT-1999;	99US-0162142.	PR	17-JUN-1999;	99US-0139452.
PR	29-OCT-1999;	99US-0162142.	PR	17-JUN-1999;	99US-0139452.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139454.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139455.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139456.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139457.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139458.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139459.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139460.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139461.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139462.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139463.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139750.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139750.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139750.
PR	29-OCT-1999;	99US-0162142.	PR	21-JUN-1999;	99US-0139817.
PR	29-OCT-1999;	99US-0162142.	PR	21-JUN-1999;	99US-0139817.
PR	29-OCT-1999;	99US-0162142.	PR	22-JUN-1999;	99US-0139899.
PR	29-OCT-1999;	99US-0162142.	PR	23-JUN-1999;	99US-0140353.
PR	29-OCT-1999;	99US-0162142.	PR	23-JUN-1999;	99US-0140354.
PR	29-OCT-1999;	99US-0162142.	PR	24-JUN-1999;	99US-0140695.
PR	29-OCT-1999;	99US-0162142.	PR	24-JUN-1999;	99US-0140695.
PR	29-OCT-1999;	99US-0162142.	PR	28-JUN-1999;	99US-0140823.
PR	29-OCT-1999;	99US-0162142.	PR	28-JUN-1999;	99US-0140823.
PR	29-OCT-1999;	99US-0162142.	PR	29-JUN-1999;	99US-0140991.
PR	29-OCT-1999;	99US-0162142.	PR	30-JUN-1999;	99US-0141287.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUL-1999;	99US-0141842.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUL-1999;	99US-0142154.
PR	29-OCT-1999;	99US-0162142.	PR	02-JUL-1999;	99US-0142055.
PR	29-OCT-1999;	99US-0162142.	PR	06-JUL-1999;	99US-0142390.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUL-1999;	99US-0142803.
PR	29-OCT-1999;	99US-0162142.	PR	09-JUL-1999;	99US-0142920.
PR	29-OCT-1999;	99US-0162142.	PR	12-JUL-1999;	99US-0142977.
PR	29-OCT-1999;	99US-0162142.	PR	13-JUL-1999;	99US-0143542.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUL-1999;	99US-0143624.
PR	29-OCT-1999;	99US-0162142.	PR	15-JUL-1999;	99US-0144005.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUL-1999;	99US-0144085.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUL-1999;	99US-0144086.
PR	29-OCT-1999;	99US-0162142.	PR	19-JUL-1999;	99US-0144325.
PR	29-OCT-1999;	99US-0162142.	PR	19-JUL-1999;	99US-0144331.
PR	29-OCT-1999;	99US-0162142.	PR	19-JUL-1999;	99US-0144332.
PR	29-OCT-1999;	99US-0162142.	PR	19-JUL-1999;	99US-0144333.
PR	29-OCT-1999;	99US-0162142.	PR	19-JUL-1999;	99US-0144334.
PR	29-OCT-1999;	99US-0162142.	PR	19-JUL-1999;	99US-0144335.
PR	29-OCT-1999;	99US-0162142.	PR	20-JUL-1999;	99US-0144352.
PR	29-OCT-1999;	99US-0162142.	PR	20-JUL-1999;	99US-0144632.

Query Match 70.5%; Score 31; DB 21; Length 99;  
Best Local Similarity 55.6%; Pred. No. 86;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATLETEVY 9  
Db 57 ASLEETQIY 65  
RESULT 13  
AAG25795  
ID AAG25795 standard; Protein; 103 AA.  
XX AC AAG25795;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 30003.  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30003.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.



```
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157863.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 70.5%; Score 31; DB 21; Length 103;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLETEVY 9
Db 61 ASLEEQIY 69

RESULT 14
ABB89417
ID ABB89417 standard; Protein; 113 AA.
XX
AC ABB89417;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 1793.
XX
KW Cytostatic; Immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
WPI; 2002-122018/16.
DR N-PSDB; ABL89826.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
disorders -
XX
PS Claim 11; SEQ ID NO 1793; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
(ABB89040-ABB90444) useful for preventing, treating or ameliorating
```

CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
 SQ Sequence 113 AA;

Query Match 70.5%; Score 31; DB 23; Length 113;

Best Local Similarity 85.7%; Pred. No. 99;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LERTEVY 9

|||||

Db 19 LERTEVF 25

RESULT 15

AAG04807  
 ID AAG04807 standard; Protein: 128 AA.

XX AAG04807;

DT 17-OCR-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 974.

KW Protein identification: signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

PN EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135153.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0138847.

PR 16-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 70.5%; Score 31; DB 21; Length 128;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATLERTEVY 9  
|:|:| |:|  
Db 57 ASLEETQIY 65

Search completed: February 6, 2003, 14:28:44  
Job time : 37 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: February 6, 2003, 14:21:37 ; Search time 15 seconds  
(without alignments)  
17.654 Million cell updates/sec

Title: us-09-641-528a-31040  
Perfect score: 44  
Sequence: 1 ATLETEVY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	70.5	260	4	US-09-589-927-2
2	31	70.5	260	4	US-09-277-665-2
3	31	70.5	448	4	US-09-134-001C-3055
4	31	70.5	457	1	US-08-206-006-2
5	31	70.5	683	4	US-09-134-001C-5576
6	30	68.2	485	4	US-09-480-921B-4
7	30	68.2	1162	4	US-09-134-001C-4008
8	30	68.2	1194	2	US-08-680-326-35
9	29	65.9	186	5	PCT-US91-08177-15
10	29	65.9	316	1	US-07-805-567-2
11	29	65.9	500	6	5171684-2
12	29	65.9	1235	2	US-08-680-326-36
13	28	63.6	301	4	US-09-586-935-5
14	28	63.6	349	4	US-09-459-774-2
15	28	63.6	389	4	US-09-311-626B-18
16	28	63.6	635	4	US-09-142-623-1
17	28	63.6	899	4	US-09-413-814-5
18	28	63.6	1009	2	US-08-680-326-31
19	28	63.6	2763	3	US-08-496-944-2
20	27	61.4	12	4	US-09-439-897-54
21	27	61.4	35	4	US-09-439-897-64
22	27	61.4	189	4	US-09-134-001C-3160
23	27	61.4	195	2	US-08-928-692-27
24	27	61.4	195	4	US-09-339-972-27
25	27	61.4	222	3	US-09-010-809-9
26	27	61.4	229	4	US-09-248-528-11
27	27	61.4	229	4	US-09-549-108-11

28	27	61.4	229	4	US-09-549-111-11	Sequence 11, Appl
29	27	61.4	229	4	US-09-549-106-11	Sequence 11, Appl
30	27	61.4	229	4	US-09-550-394-11	Sequence 11, Appl
31	27	61.4	385	2	US-08-387-942C-8	Sequence 8, Appl
32	27	61.4	385	2	US-08-387-942C-24	Sequence 24, Appl
33	27	61.4	385	2	US-08-387-942C-26	Sequence 26, Appl
34	27	61.4	385	2	US-08-387-942C-27	Sequence 27, Appl
35	27	61.4	470	2	US-08-959-011-1	Sequence 1, Appl
36	27	61.4	496	4	US-08-924-183-10	Sequence 10, Appl
37	27	61.4	496	4	US-09-488-364-10	Sequence 10, Appl
38	27	61.4	496	4	US-09-113-785-3	Sequence 3, Appl
39	27	61.4	872	2	US-08-387-942C-5	Sequence 5, Appl
40	27	61.4	997	2	US-08-387-942C-4	Sequence 4, Appl
41	27	61.4	1389	2	US-08-619-198-5	Sequence 5, Appl
42	27	61.4	1403	2	US-08-387-942C-3	Sequence 3, Appl
43	27	61.4	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
44	27	61.4	5087	4	US-09-144-085-1	Sequence 1, Appl
45	27	61.4	6095	4	US-09-144-085-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-589-927-2  
; Sequence 2, Application US/09589927  
; Patent No. 6432706  
; GENERAL INFORMATION:  
; APPLICANT: University of Kansas Medical Center  
; TITLE OF INVENTION: The use of isolated domains of Type IV Collagen to  
; FILE OF INVENTION: Modify Cell and Tissue Interactions  
; FILE REFERENCE: 945251  
; CURRENT APPLICATION NUMBER: US/09/589,927  
; CURRENT FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Human  
US-09-589-927-2

Query Match 70.5%; Score 31; DB 4; Length 260;  
Best Local Similarity 55.6%; Pred. No. 52;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLETEVY 9  
Db 225 ATIERSEMF 233

RESULT 2  
US-09-277-665-2  
; Sequence 2, Application US/09277665  
; Patent No. 6440729  
; GENERAL INFORMATION:  
; APPLICANT: University of Kansas Medical Center  
; TITLE OF INVENTION: The use of isolated domains of Type IV Collagen to  
; FILE OF INVENTION: Modify Cell and Tissue Interactions  
; FILE REFERENCE: 94525-1  
; CURRENT APPLICATION NUMBER: US/09/277,665  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Human  
US-09-277-665-2

Query Match 70.5%; Score 31; DB 4; Length 260;  
Best Local Similarity 55.6%; Pred. No. 52;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
; MOLECULE TYPE: protein
; US-08-206-006-2

Query Match      70.5%; Score 31; DB 1; Length 457;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLRTVEY 9
Db 351 TIERTQVF 358
      |::|::|:

RESULT 5
US-09-134-001C-5576
; Sequence 5576, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5576
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5576

Query Match      70.5%; Score 31; DB 4; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ERTEYV 9
Db 534 ERTEYV 539
      |::|::|:

RESULT 6
US-09-480-921B-4
; Sequence 4, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Wegrich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/S-30780A
; CURRENT APPLICATION NUMBER: US/09/480,921B
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-480-921B-4

Query Match      68.2%; Score 30; DB 4; Length 485;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEYV 9
Db 275 LEQTEYV 281
      |::|::|:

; MOLECULE TYPE: protein
; US-08-206-006-2

Query Match      70.5%; Score 31; DB 1; Length 457;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLRTVEY 9
Db 351 TIERTQVF 358
      |::|::|:

RESULT 5
US-09-134-001C-5576
; Sequence 5576, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3055
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3055

Query Match      70.5%; Score 31; DB 4; Length 448;
Best Local Similarity 66.7%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATLRTVEY 9
Db 48 STCNRTVEY 56
      |::|::|:

RESULT 4
US-08-206-006-2
; Sequence 2, Application US/08206006
; Patent No. 5474912
; GENERAL INFORMATION:
; APPLICANT: Sherman, David H.
; APPLICANT: Hu, Wei-Shou
; APPLICANT: Malmberg, Li-Hong
; TITLE OF INVENTION: Method for Increasing Production of
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lorraine R. Sherman
; STREET: 3201 Cavell Lane
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: U.S.A.
; ZIP: 55426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,006
; FILING DATE: 03-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherman, Lorraine R.
; REGISTRATION NUMBER: 30,105
; REFERENCE/DOCKET NUMBER: 6102
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
RESULT 7
US-09-134-001C-4008
; Sequence 4008, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4008
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4008
Query Match 68.2%; Score 30; DB 4; Length 1162;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
   |||: |||
Db 956 ATLDLTKVY 964

RESULT 8
US-08-680-326-35
; Sequence 35, Application US/08680326
; Patent No. 5925733
; GENERAL INFORMATION:
; APPLICANT: ROSE, TIMOTHY M.
; APPLICANT: BOSCH, MARNIX
; APPLICANT: STRAND, KURT
; APPLICANT: TODARO, GEORGE J.
; TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
; TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
; NUMBER OF SEQUENCES: 152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,326
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Schliff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20001.00
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
US-08-680-326-35
Query Match 68.2%; Score 30; DB 2; Length 1194;
Best Local Similarity 71.4%; Pred. No. 4.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
   |||: |||
Db 234 LERTDVI 240

RESULT 9
PCT-US91-08177-15
; Sequence 15, Application PC/TUS9108177
; GENERAL INFORMATION:
; APPLICANT: Samal, Siba K
; TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08177
; FILING DATE: 19911104
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/608,937
; FILING DATE: 05-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hightet, David W.
; REGISTRATION NUMBER: 30,265
; REFERENCE/DOCKET NUMBER: 20509-96711
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4854
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 186 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-08177-15
Query Match 65.9%; Score 29; DB 5; Length 186;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
   |||: |||
Db 64 AELDRTEVY 72

RESULT 10
US-07-805-567-2
; Sequence 2, Application US/07805567
; Patent No. 5378457
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Tartaglia, James
; TITLE OF INVENTION: INTERFERON SENSITIVE RECOMBINANT
; TITLE OF INVENTION: POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Curtis, Morris & Safford  
ADDRESSEE: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/805,567  
FILING DATE: 19911216  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/638,080  
FILING DATE: 07-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2281  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 316 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-805-567-2

Query Match 65.9%; Score 29; DB 1; Length 316;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTE 7  
| | | | |  
DB 242 TLERTE 247

RESULT 11  
5171684-2  
PATENT NO. 5171684  
APPLICANT: YEN, KWANG-MU; BLATT, LAWRENCE M.; KARL, MICHAEL R.  
TITLE OF INVENTION: BIOCONVERSIONS CATALYZED BY THE TOULUENE  
MONOOXYGENASE OF PSEUDOMONAS MENDOCINA KR-1  
NUMBER OF SEQUENCES: 41  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/590,374  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 177,631  
FILING DATE: 05-APR-1988  
SEQ ID NO: 2:  
LENGTH: 500  
5171684-2

Query Match 65.9%; Score 29; DB 6; Length 500;  
Best Local Similarity 55.6%; Pred. No. 2.8e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
| | | | |  
DB 74 AALERAKIY 82

RESULT 12  
US-08-680-326-36  
Sequence 36, Application US/08680326  
Patent No. 5925733

GENERAL INFORMATION:  
APPLICANT: ROSE, TIMOTHY M.  
APPLICANT: BOSCH, MARNIX  
APPLICANT: STRAND, KURT  
APPLICANT: TODARO, GEORGE J.  
TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES  
TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL  
FIBROMATOSIS  
NUMBER OF SEQUENCES: 152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/680,326  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Schiff, J. Michael  
REGISTRATION NUMBER: 40,253  
REFERENCE/DOCKET NUMBER: 29938-20001.00  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1235 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-680-326-36

Query Match 65.9%; Score 29; DB 2; Length 1235;  
Best Local Similarity 71.4%; Pred. No. 7.8e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEVY 9  
| | | | |  
DB 253 VERTDVI 259

RESULT 13  
US-09-586-935-5  
Sequence 5, Application US/09586935  
Patent No. 6191267  
GENERAL INFORMATION:  
APPLICANT: KONG, HUIMIN  
APPLICANT: HIGGINS, LAUREN S.  
APPLICANT: DALTON, MICHAEL  
APPLICANT: KUCERA, REBECCA B.  
APPLICANT: SCHILDKRAUT, IRA  
TITLE OF INVENTION: Cloning And Producing The N.BstNBI Nicking Endonuclease  
FILE REFERENCE: NEB-178  
CURRENT APPLICATION NUMBER: US/09/586,935  
CURRENT FILING DATE: 2000-06-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 301  
TYPE: PRT  
ORGANISM: Bacillus stearothermophilus  
US-09-586-935-5

Query Match 63.6%; Score 28; DB 4; Length 301;



Search completed: February 6, 2003, 14:27:19  
Job time : 16 secs

Best Local Similarity 57.1%; Pred. NO. 2.6e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEVY 9  
|:||||:|  
Db 173 LOKTEIY 179

RESULT 14  
US-09-459-774-2  
; Sequence 2, Application US/09459774  
; Patent No. 6297030  
; GENERAL INFORMATION:  
; APPLICANT: Michael Robert Barnes  
; APPLICANT: Tania Tamson Testa  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP-30193  
; CURRENT APPLICATION NUMBER: US/09/459,774  
; CURRENT FILING DATE: 1999-12-13  
; EARLIER APPLICATION NUMBER: UK 9828419.3  
; EARLIER FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 349  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-459-774-2

Query Match 63.6%; Score 28; DB 4; Length 349;  
Best Local Similarity 83.3%; Pred. NO. 3.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 ERTEVY 9  
||||:|  
Db 341 ERTEMY 346

RESULT 15  
US-09-311-626B-18  
; Sequence 18, Application US/09311626B  
; Patent No. 6399347  
; GENERAL INFORMATION:  
; APPLICANT: Jorgensen, Per Lina  
; APPLICANT: Schnort, Kirk  
; APPLICANT: Andersen, Lene No. 6399347boe  
; APPLICANT: Schulein, Martin  
; APPLICANT: Outtrup, Helle  
; TITLE OF INVENTION: NO. 6399347el Rhamnogalacturonan Hydrolases  
; FILE REFERENCE: 5572.204-US  
; CURRENT APPLICATION NUMBER: US/09/311,626B  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: 0608/98  
; PRIOR FILING DATE: 1998-05-01  
; PRIOR APPLICATION NUMBER: 60/084,358  
; PRIOR FILING DATE: 1998-05-05  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Bacillus halodurans KJ59  
US-09-311-626B-18

Query Match 63.6%; Score 28; DB 4; Length 389;  
Best Local Similarity 85.7%; Pred. NO. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
|||||  
Db 155 TLERVEV 161

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:27:47 ; Search time 11 Seconds  
(without alignments)  
18.139 Million cell updates/sec

Title: US-09-641-528A-31040  
Perfect score: 44  
Sequence: 1 ATLETERVEY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues  
Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	72.7	159	10	US-09-815-242-5454
2	32	72.7	162	10	US-09-815-242-12683
3	32	72.7	383	9	US-10-224-446-2
4	32	72.7	1242	10	US-09-925-299-911
5	31	70.5	406	10	US-09-925-302-507
6	30	68.2	255	9	US-10-224-446-28
7	30	68.2	255	9	US-10-224-446-30
8	30	68.2	485	9	US-10-047-412A-4
9	30	68.2	514	10	US-09-886-468-23
10	30	68.2	548	9	US-09-738-626-3980
11	30	68.2	1127	10	US-09-815-242-5373
12	30	68.2	1158	10	US-09-815-242-12522
13	29	65.9	9	1	US-08-821-739A-38
14	29	65.9	59	10	US-09-864-761-44586
15	29	65.9	68	10	US-09-864-761-40150
16	29	65.9	68	10	US-09-864-761-47126
17	29	65.9	268	10	US-09-925-301-1315
18	29	65.9	424	10	US-09-815-242-12587
19	29	65.9	425	10	US-09-815-242-5485

20	29	65.9	1235	10	US-09-904-065-6	Sequence 6, Appli
21	29	65.9	1235	10	US-09-904-065-8	Sequence 8, Appli
22	29	65.9	1235	10	US-09-904-065-10	Sequence 10, Appli
23	29	65.9	1235	10	US-09-904-065-16	Sequence 16, Appli
24	29	65.9	1235	10	US-09-904-065-17	Sequence 17, Appli
25	29	65.9	1235	10	US-09-904-065-18	Sequence 18, Appli
26	29	65.9	1235	10	US-09-904-065-19	Sequence 19, Appli
27	28	63.6	345	10	US-09-815-242-5854	Sequence 5854, Ap
28	28	63.6	349	10	US-09-903-817-2	Sequence 2, Appli
29	28	63.6	358	10	US-09-815-242-12866	Sequence 12866, A
30	28	63.6	383	9	US-10-042-431-40	Sequence 40, Appli
31	28	63.6	383	9	US-09-759-130B-410	Sequence 410, App
32	28	63.6	440	10	US-09-815-242-11399	Sequence 11399, A
33	28	63.6	453	9	US-10-114-893-184	Sequence 184, App
34	28	63.6	453	9	US-09-909-567B-55	Sequence 55, Appli
35	28	63.6	453	10	US-09-529-063-61	Sequence 61, Appli
36	28	63.6	472	9	US-09-738-626-6752	Sequence 6752, Ap
37	28	63.6	530	9	US-09-738-626-4780	Sequence 4780, Ap
38	28	63.6	626	10	US-09-862-027-43	Sequence 43, Appli
39	28	63.6	628	10	US-09-862-027-48	Sequence 48, Appli
40	28	63.6	635	9	US-09-990-385-1	Sequence 1, Appli
41	27	61.4	59	9	US-09-796-692-1454	Sequence 1454, Ap
42	27	61.4	59	9	US-09-796-692-1904	Sequence 1904, Ap
43	27	61.4	130	9	US-09-738-626-5055	Sequence 5055, Ap
44	27	61.4	140	9	US-09-738-626-5368	Sequence 5368, Ap
45	27	61.4	178	9	US-09-764-868-1036	Sequence 1036, Ap

ALIGNMENTS

RESULT 1  
US-09-815-242-5454  
; Sequence 5454, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA-011A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5454  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-5454  
Query Match 72.7%; Score 32; DB 10; Length 159;  
Best Local Similarity 75.0%; Pred. No. 8.1;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8  
||:|:|  
Db 95 ATMEKTEV 102

## RESULT 2

US-09-815-242-12683  
; Sequence 12683, Application US/09815242  
; Patent No. US20020061569A1

## GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA 011A

; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 12683

; LENGTH: 162

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12683

Query Match 72.7%; Score 32; DB 10; Length 162;

Best Local Similarity 75.0%; Pred. No. 8.2;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8  
||:|:|  
Db 98 ATMEKTEV 105

## RESULT 3

US-10-224-446-2

; Sequence 2, Application US/1022446

; Publication No. US20030024010A1

## GENERAL INFORMATION:

; APPLICANT: CAHOON, EDGAR B.

; APPLICANT: CARLSON, THOMAS J.

; APPLICANT: HITZ, WILLIAM D.

; APPLICANT: RIPP, KEVIN G.

; TITLE OF INVENTION: GENES FOR PLANT FATTY ACID MODIFYING ENZYMES

; TITLE OF INVENTION: ASSOCIATED

; TITLE OF INVENTION: WITH CONJUGATED DOUBLE BOND FORMATION

; FILE REFERENCE: BB-1249

; CURRENT APPLICATION NUMBER: US/10/224,446

; CURRENT FILING DATE: 2002-08-20

; PRIOR APPLICATION NUMBER: US/09/375,299

; PRIOR FILING DATE: 1999-08-16

; PRIOR APPLICATION NUMBER: 60/097,186  
; PRIOR FILING DATE: August 20, 1998  
; PRIOR APPLICATION NUMBER: 60/142,756  
; PRIOR FILING DATE: July 8, 1999  
; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: MICROSOFT OFFICE 97

; SEQ ID NO 2

; LENGTH: 383

; TYPE: PRT

; ORGANISM: Impatiens balsamina

US-10-224-446-2

Query Match 72.7%; Score 32; DB 9; Length 383;

Best Local Similarity 66.7%; Pred. No. 21;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEV 9  
||:|:|  
Db 149 ASIERDEVY 157

## RESULT 4

US-09-925-299-911

; Sequence 911, Application US/09925299

; Patent No. US20020055627A1

## GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA102

; CURRENT APPLICATION NUMBER: US/09/925,299

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05883

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 911

; LENGTH: 1242

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (224)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (1013)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (1034)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-299-911

Query Match 72.7%; Score 32; DB 10; Length 1242;

Best Local Similarity 66.7%; Pred. No. 79;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEV 9  
||:|:|  
Db 393 ATQERTDLY 401

## RESULT 5

US-09-925-302-507

; Sequence 507, Application US/09925302

; Patent No. US20020044941A1

## GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA104

; CURRENT APPLICATION NUMBER: US/09/925,302

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05918

; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 507  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (71)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-302-507

Query Match 70.5%; Score 31; DB 10; Length 406;  
Best Local Similarity 55.6%; Pred. No. 37;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
||:||||:  
Db 371 ATIERSEMF 379

## RESULT 6

US-10-224-446-28  
; Sequence 28, Application US/10224446  
; Publication No. US20030024010A1  
; GENERAL INFORMATION:  
; APPLICANT: CAHOON, EDGAR B.  
; APPLICANT: CARLSON, THOMAS J.  
; APPLICANT: HITZ, WILLIAM D.  
; APPLICANT: RIPP, KEVIN G.  
; TITLE OF INVENTION: GENES FOR PLANT FATTY ACID MODIFYING ENZYMES  
; TITLE OF INVENTION: WITH CONJUGATED DOUBLE BOND FORMATION  
; FILE REFERENCE: BB-1249  
; CURRENT APPLICATION NUMBER: US/10/224,446  
; CURRENT FILING DATE: 2002-08-20  
; PRIOR APPLICATION NUMBER: US/09/375,299  
; PRIOR FILING DATE: 1999-08-16  
; PRIOR APPLICATION NUMBER: 60/097,186  
; PRIOR FILING DATE: August 20, 1998  
; PRIOR APPLICATION NUMBER: 60/142,756  
; PRIOR FILING DATE: July 8, 1999  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 28  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Aleurites fordii  
US-10-224-446-28

Query Match 68.2%; Score 30; DB 9; Length 255;  
Best Local Similarity 83.3%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 ERTEVY 9  
||||:|  
Db 184 ERTEY 189

## RESULT 7

US-10-224-446-30  
; Sequence 30, Application US/10224446  
; Publication No. US20030024010A1  
; GENERAL INFORMATION:  
; APPLICANT: CAHOON, EDGAR B.  
; APPLICANT: CARLSON, THOMAS J.  
; APPLICANT: HITZ, WILLIAM D.  
; APPLICANT: RIPP, KEVIN G.  
; TITLE OF INVENTION: GENES FOR PLANT FATTY ACID MODIFYING ENZYMES  
; TITLE OF INVENTION: ASSOCIATED

; TITLE OF INVENTION: WITH CONJUGATED DOUBLE BOND FORMATION  
; FILE REFERENCE: BB-1249  
; CURRENT APPLICATION NUMBER: US/10/224,446  
; CURRENT FILING DATE: 2002-08-20  
; PRIOR APPLICATION NUMBER: US/09/375,299  
; PRIOR FILING DATE: 1999-08-16  
; PRIOR APPLICATION NUMBER: 60/097,186  
; PRIOR FILING DATE: August 20, 1998  
; PRIOR APPLICATION NUMBER: 60/142,756  
; PRIOR FILING DATE: July 8, 1999  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: MICROSOFT OFFICE 97  
; SEQ ID NO 30  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Aleurites fordii  
US-10-224-446-30

Query Match 68.2%; Score 30; DB 9; Length 255;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
|:|||||:  
Db 112 ASLDERVEF 120

## RESULT 8

US-10-047-412A-4  
; Sequence 4, Application US/10047412A  
; Publication No. US20020197696A1  
; GENERAL INFORMATION:  
; APPLICANT: Levin, Joshua Z.  
; APPLICANT: Budziszewski, Gregory J.  
; APPLICANT: Potter, Sharon L.  
; APPLICANT: Wegrich, Lynette M.  
; TITLE OF INVENTION: Herbicide Target Genes and Methods  
; FILE REFERENCE: PB/5-30780DIV  
; CURRENT APPLICATION NUMBER: US/10/047,412A  
; CURRENT FILING DATE: 2002-04-11  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-047-412A-4

Query Match 68.2%; Score 30; DB 9; Length 485;  
Best Local Similarity 71.4%; Pred. No. 74;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEVY 9  
||:||||:  
Db 275 LEQTEY 281

## RESULT 9

US-09-886-468-23  
; Sequence 23, Application US/09886468  
; Patent No. US20020037293A1  
; GENERAL INFORMATION:  
; APPLICANT: Aventis Pasteur Limited  
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses th  
; FILE REFERENCE: 77813-5  
; CURRENT APPLICATION NUMBER: US/09/886,468  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/113,280  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,281  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,282  
; PRIOR FILING DATE: 1998-12-23

QY 2 TLERTEV 9  
|:|:|:|:  
Db 89 TLDRTDVF 96

RESULT 11  
US-09-815-242-5373  
; Sequence 5373, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 5373  
; LENGTH: 1127  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-5373

Query Match 68.2%; Score 30; DB 10; Length 514;  
Best Local Similarity 85.7%; Pred. No. 79;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
|:|:|:|:  
Db 461 TLERTQV 467

RESULT 10  
US-09-738-626-3980  
; Sequence 3980, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, TOSHIO  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3980  
; LENGTH: 548  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3980

Query Match 68.2%; Score 30; DB 9; Length 548;  
Best Local Similarity 62.5%; Pred. No. 85;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Query Match 68.2%; Score 30; DB 10; Length 1127;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLTERTEV 9  
|:|:|:|:  
Db 947 ATLDLTKVY 955

RESULT 12  
US-09-815-242-12522  
; Sequence 12522, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12522  
; LENGTH: 1158  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-12522

Query Match 68.2%; Score 30; DB 10; Length 1158;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
|||: |||  
Db 953 ATLDLTKVY 961

RESULT 13  
US-08-821-739A-38  
; Sequence 38, Application US/08821739A  
; Patent No. US20020168374A1  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Celis, Esteban  
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses  
; FILE REFERENCE: 2060.005000A  
; CURRENT APPLICATION NUMBER: US/08/821.739A  
; CURRENT FILING DATE: 1999-03-20  
; PRIOR APPLICATION NUMBER: 60/013,833  
; PRIOR FILING DATE: 1996-03-21  
; PRIOR APPLICATION NUMBER: 08/589,107  
; PRIOR FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 08/451,913  
; PRIOR FILING DATE: 1995-05-26  
; PRIOR APPLICATION NUMBER: 08/347,610  
; PRIOR FILING DATE: 1994-12-01  
; PRIOR APPLICATION NUMBER: 08/186,266  
; PRIOR FILING DATE: 1994-01-25  
; PRIOR APPLICATION NUMBER: 08/159,339  
; PRIOR FILING DATE: 1993-11-29  
; PRIOR APPLICATION NUMBER: 08/103,396  
; PRIOR FILING DATE: 1993-08-06  
; PRIOR APPLICATION NUMBER: 08/027,746  
; PRIOR FILING DATE: 1993-03-05  
; PRIOR APPLICATION NUMBER: 07/926,666  
; PRIOR FILING DATE: 1992-08-07  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 38  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-821-739A-38

Query Match 65.9%; Score 29; DB 1; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.1e+05;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTEVY 9  
:|||: |||  
Db 2 VERTDYY 8

RESULT 14  
US-09-864-761-44586  
; Sequence 44586, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864.761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 44586  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO ALL33377.4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.77  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.83  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.61  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.66  
; OTHER INFORMATION: EST\_HUMAN HIT: AA368914.1, EVALUATE 1.60e+00  
; OTHER INFORMATION: SWISSPROT HIT: P08548, EVALUATE 2.70e+00  
US-09-864-761-44586

Query Match 65.9%; Score 29; DB 10; Length 59;

Best Local Similarity 55.6%; Pred. No. 12; Mismatches 2; Indels 2; Gaps 0;  
Matches 5; Conservative 2; Mismatches 2; Indels 2; Gaps 0;  
Qy 1 ATLERTEVY 9  
Db 47 STLRTQTY 55

Query Match 65.9%; Score 29; DB 10; Length 68;  
Best Local Similarity 85.7%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATLERTE 7  
Db 39 ATLEQTE 45

Search completed: February 6, 2003, 14:29:58  
Job time : 12 secs

RESULT 15  
US-09-864-761-40150  
; Sequence 40150, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 40150  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005323.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.91  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
; OTHER INFORMATION: EST\_HUMAN HIT: AW950723.1, EVALUATE 2.00e-30  
; OTHER INFORMATION: SWISSPROT HIT: P12882, EVALUATE 4.00e-31  
US-09-864-761-40150



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:27:27 ; Search time 16 Seconds  
(without alignments)  
54.076 Million cell updates/sec

Title: US-09-641-528A-31040  
Perfect score: 44  
Sequence: 1 ATLETERVEY 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	44	100.0	158	2 S36561	E6 protein - human
2	34	77.3	350	2 F75259	glutaryl-tRNA redu
3	33	75.0	149	1 W6WL58	E6 protein - human
4	33	75.0	1264	2 H96498	hypothetical prote
5	32	72.7	90	2 A82562	hypothetical prote
6	32	72.7	90	2 E82562	hypothetical prote
7	32	72.7	158	1 W6WLPR	E6 protein - human
8	32	72.7	215	2 C82493	hypothetical prote
9	32	72.7	368	2 T20182	hypothetical prote
10	32	72.7	662	2 G95143	exonuclease ABC,
11	32	72.7	662	2 E98011	exonuclease ABC ch
12	32	72.7	668	2 A42385	exonuclease ABC c
13	32	72.7	692	2 B86695	exonuclease ABC s
14	32	72.7	1462	1 AG2292	protein-tyrosine-p
15	31	70.5	167	2 AG2292	hypothetical prote
16	31	70.5	183	2 S75697	hypothetical prote
17	31	70.5	245	2 C83812	hypothetical prote
18	31	70.5	258	2 B61228	hypothetical prote
19	31	70.5	382	2 T22798	collagen alpha 1(I
20	31	70.5	448	2 F89950	hypothetical prote
21	31	70.5	457	2 A38169	lysine aminotransf
22	31	70.5	582	2 C90165	conserved hypotet
23	31	70.5	1223	2 S62011	PHO85 protein - ye
24	31	70.5	1451	2 S42167	190K protein - hum
25	31	70.5	1669	1 CGH04B	collagen alpha 1(I
26	31	70.5	1669	1 CGMS4B	collagen alpha 1(I
27	31	70.5	1877	2 T21861	hypothetical prote
28	30	68.2	58	2 S48098	cryptic protein -
29	30	68.2	90	2 C82562	hypothetical prote

30	30	68.2	114	1 C72316	aspartate 1-decarb
31	30	68.2	123	2 F70202	hypothetical prote
32	30	68.2	168	2 T23105	hypothetical prote
33	30	68.2	253	2 T35506	hypothetical prote
34	30	68.2	259	2 H71302	probable ABC trans
35	30	68.2	294	2 T22145	hypothetical prote
36	30	68.2	329	2 D95372	probable integrase
37	30	68.2	329	2 D95409	probable integrase
38	30	68.2	365	2 C69538	hypothetical prote
39	30	68.2	372	2 G97844	hypothetical prote
40	30	68.2	384	2 E81689	conserved hypotet
41	30	68.2	399	2 T10314	viral transcriptio
42	30	68.2	419	2 C82109	glutaryl-tRNA redu
43	30	68.2	425	2 T29617	hypothetical prote
44	30	68.2	429	2 F72784	probable dihydroor
45	30	68.2	432	2 C84351	histidyl-tRNA synt

ALIGNMENTS

RESULT 1  
S36561  
E6 protein - human papillomavirus type 45  
C:Species: human papillomavirus type 45  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S36561  
R:Delius, H.; Hofmann, B.  
submitted to the EMBL Data Library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36561  
A:Molecule type: DNA  
A:Residues: 1-158 <DEL>  
A:Cross-references: EMBL:X74479; NID:g397022; PIDN:CAA52573.1; PID:g397023  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 100.0%; Score 44; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLETERVEY 9  
Db 37 ATLETERVEY 45

RESULT 2  
F75259  
glutaryl-tRNA reductase - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: F75259  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: F75259  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-350 <WHI>  
A:Cross-references: GB:AE002084; GB:AE000513; NID:g6460372; PIDN:AAF12088.1; PID:g646

Query Match 77.3%; Score 34; DB 2; Length 350;  
Best Local Similarity 77.8%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1  
Db 1

```
Qy 1 ATLERTEVY 9
||| |||||
Db 52 ATCNRTEVY 60

RESULT 3
W6WL58
E6 protein - human papillomavirus type 58
C:Species: human papillomavirus type 58
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: E36779
R:Kiril, Y.; Iwamoto, S.; Matsukura, T.
Virology 185, 424-427, 1991
A:Title: Human papillomavirus type 58 DNA sequence.
A:Reference number: A36779; MUID:92024102; PMID:1656594
A:Accession: E36779
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-149 <KIR>
A:Cross-references: GB:D90400; NID:G222386; PIDN:BAA31845.1; PID:G3337098
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; zinc finger
F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif

Query Match 75.0%; Score 33; DB 1; Length 149;
Best Local Similarity 75.0%; Pred. No. 9.7;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
||| |||||
Db 36 TLORSEVY 43

RESULT 4
H96498
hypotheical protein T10P12.3 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96498
R:theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maftl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1264 <STO>
A:Cross-references: GB:AE005173; NID:G5080762; PIDN:AAD39272.1; GSPDB:GN00141
C:Genetics:
A:Gene: T10P12.3
A:Map position: 1

Query Match 75.0%; Score 33; DB 2; Length 1264;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8
:|||||
Db 62 STLERTEI 69

RESULT 5
A82562
hypotheical protein XF2401 [imported] - Xylella fastidiosa (strain 9a5c)
```

```
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: A82562
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <SIM>
A:Cross-references: GB:AE004049; GB:AE003849; NID:G9107579; PIDN:AAF85200.1; GSPDB:GN
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2401

Query Match 72.7%; Score 32; DB 2; Length 90;
Best Local Similarity 66.7%; Pred. No. 9.3;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
||| |||||
Db 12 ATLDRAEAY 20

RESULT 6
E82562
hypotheical protein XF2405 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82562
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <SIM>
A:Cross-references: GB:AE004049; GB:AE003849; NID:G9107579; PIDN:AAF85204.1; GSPDB:GN
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2405
```

hypoetical protein C53B4.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000  
C:Accession: T20182  
R:Berks, M.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: Z19233  
A:Accession: T20182  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-368 <WIL>  
A:Cross-references: EMBL:Z68215; PIDN:CAA92455.2; GSPDB:GN00022; CESP:C53B4.7  
A:Experimental source: clone C53B4  
C:Genetics:  
A:Gene: CESP:C53B4.7  
A:Map position: 4  
A:Introns: 30/3; 101/3; 212/1; 348/3  
C:Superfamily: GDP-D-mannose dehydratase

Query Match 72.7%; Score 32; DB 2; Length 368;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
: : : : :  
Db 95 STIEPTEVY 103

RESULT 10  
G95143  
excinuclease ABC, chain B [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
C:Accession: G95143  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H  
nson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: G95143  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-662 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK75344.1; PID:gl4972720; GSPDB:GN00164; TIGR:  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SPI238  
C:Superfamily: excinuclease ABC chain B

Query Match 72.7%; Score 32; DB 2; Length 662;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEV 8  
: : : : :  
Db 486 TLERTEI 492

RESULT 11  
E98011  
exonuclease ABC chain B [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
C:Accession: E98011  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B. S.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234

Query Match 72.7%; Score 32; DB 2; Length 90;  
Best Local Similarity 66.7%; Pred. No. 9.3;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
: : : : :  
Db 12 ATLDRAEAY 20

RESULT 7  
W6WLR  
E6 protein - human papillomavirus type ME180 (provirus)  
C:Species: human papillomavirus type ME180  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Feb-1997  
C:Accession: C40509  
R:Reuter, S.; Delius, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwarz, E.  
J. Virol. 65, 5564-5568, 1991  
A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma  
A:Reference number: A40509; MUID:91374616; PMID:1716694  
A:Accession: C40509  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-158 <REG>  
A:Cross-references: GB:M73258  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; zinc finger  
F:32-68/Region: zinc finger CCCC motif  
F:103-141/Region: zinc finger CCCC motif

Query Match 72.7%; Score 32; DB 1; Length 158;  
Best Local Similarity 85.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps - 0;

Qy 3 LERTEVY 9  
: : : : :  
Db 39 LQTEVY 45

RESULT 8  
CB2493  
hypoetical protein VCA0170 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: CB2493  
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: CB2493  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-215 <HEI>  
A:Cross-references: GB:AE004357; GB:AE003853; NID:G9657547; PIDN:AAF96083.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA0170  
A:Map position: 2

Query Match 72.7%; Score 32; DB 2; Length 215;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTEVY 9  
: : : : :  
Db 144 LERTELY 150

RESULT 9  
T20182

A:Accession: E98011  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-662 <KUR>  
A:Cross-references: GB:AF007317; PIDN:AAK99921.1; PID:g15458744; GSPDB:GN00174  
C:Genetics:  
A:Gene: uvrB  
C:Superfamily: excinuclease ABC chain B

Query Match 72.7%; Score 32; DB 2; Length 662;  
Best Local Similarity 85.7%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
|||||  
Db 486 TLERTEI 492

## RESULT 12

A42385  
exinuclease ABC chain B - Streptococcus pneumoniae plasmid pSB470

N:Alternate names: uvrB protein  
N:Contains: excision endonuclease ABC (EC 3.1.1.-) chain B  
C:Species: Streptococcus pneumoniae  
C:Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 02-Feb-2001  
C:Accession: A42385  
R:Sicard, N.; Oreglia, J.; Estevenson, A.M.  
J. Bacteriol. 174, 2412-2415, 1992

A:Title: Structure of the gene complementing uvr-402 in Streptococcus pneumoniae: homologous

A:Reference number: A42385; MUID:92202177; PMID:1551859

A:Accession: A42385

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-668 <SIC>

A:Cross-references: EMBL:M80215; NID:g153853; PIDN:AAK27020.1; PID:g153854

C:Genetics:

C:Superfamily: excinuclease ABC chain B

C:Keywords: ATP; DNA repair; hydrolase; nucleotide binding; P-loop

F:50-57/Region: nucleotide-binding motif A (P-loop)

F:345-350/Region: nucleotide-binding motif B

F:349-352/Region: DEXH motif

Query Match 72.7%; Score 32; DB 2; Length 668;  
Best Local Similarity 85.7%; Pred. No. 84;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
|||||  
Db 492 TLERTEI 498

## RESULT 13

B86695

exinuclease ABC subunit B [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C:Accession: B86695

R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: B86695

A:Molecule type: DNA

A:Residues: 1-692 <STO>

A:Cross-references: GB:AE005176; PID:g17273449; PIDN:AAK04660.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: uvrB

C:Superfamily: excinuclease ABC chain B

Query Match 72.7%; Score 32; DB 2; Length 692;

Best Local Similarity 85.7%; Pred. No. 87;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
|||||  
Db 487 TLERTEI 493

## RESULT 14

B36182

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type PTP69D precursor - fruit fly  
N:Alternate names: protein-tyrosine-phosphatase DPTP  
C:Species: Drosophila melanogaster

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: B36182

R:Streuli, M.; Krueger, N.X.; Tsai, A.Y.M.; Saito, H.

Proc. Natl. Acad. Sci. U.S.A. 86, 8698-8702, 1989

A:Title: A family of receptor-linked protein tyrosine phosphatases in humans and Dros

A:Reference number: A36182; MUID:90046860; PMID:2554325

A:Accession: B36182

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1462 <STR>

A:Cross-references: GB:M27699; NID:g158188; PIDN:AAA28842.1; PID:g158189

C:Genetics:

A:Gene: Ptp69D

A:Cross-references: FlyBase:FBgn0014007

C:Superfamily: protein-tyrosine-phosphatase, receptor type PTP69D; fibronectin type I  
phosphatase homology

C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembra

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-1462/Product: protein-tyrosine-phosphatase, receptor type PTP69D #status predict

F:29-806/Domain: extracellular #status predicted <EXT>

F:38-114/Domain: immunoglobulin homology <IM1>

F:147-216/Domain: immunoglobulin homology <IM2>

F:807-823/Domain: transmembrane #status predicted <TMN>

F:824-1462/Domain: intracellular #status predicted <INT>

F:917-1145/Domain: protein-tyrosine-phosphatase homology <PTP1>

F:1213-1439/Domain: protein-tyrosine-phosphatase homology <PTP2>

F:45-112,154-214/Disulfide bonds: #status predicted

F:1097/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1103/Binding site: substrate phosphate (Arg) #status predicted

F:1391/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1397/Binding site: substrate phosphate (Arg) #status predicted

Query Match 72.7%; Score 32; DB 1; Length 1462;

Best Local Similarity 87.5%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLERTEVY 9  
||| ||||  
Db 198 TLEFTEVY 205

## RESULT 15

AG2292

hypothetical protein all3894 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AG2292

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759640

A:Accession: AG2292

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-167 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAE75593.1; PID:g17133028; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all3894

Query Match 70.5%; Score 31; DB 2; Length 167;  
Best Local Similarity 75.0%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLERTEVY 9  
   | | | | |  
Db 108 TLEQTRVY 115

Search completed: February 6, 2003, 14:29:06  
Job time : 18 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:26:51 ; Search time 11 Seconds  
(without alignments)  
33.935 Million cell updates/sec

Title: US-09-641-528a-31040  
Perfect score: 44  
Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	44	100.0	158	VE6_HPV45	P21735 human papillomavirus type 45
2	34	77.3	350	HEM1_DEIRA	Q99rs5 deinoococcus
3	33	75.0	149	VE6_HPV58	P26555 human papillomavirus type 58
4	33	75.0	437	HEM1_PAEMA	O69108 paenibacillus
5	32	72.7	158	VE6_HPV68	P54667 human papillomavirus type 68
6	32	72.7	158	VE6_HPVME	P27962 human papillomavirus type 68
7	32	72.7	368	GMD1_CAEEL	Q18801 caenorhabditis
8	32	72.7	662	UVRB_STRPN	Q54986 streptococcus
9	32	72.7	663	UVRB_STRPY	Q99za5 streptococcus
10	32	72.7	692	UVRB_LACLA	Q9ci06 lactococcus
11	32	72.7	1462	PTP6_DROME	P16620 drosophila
12	31	70.5	382	GMD2_CAEEL	O45583 caenorhabditis
13	31	70.5	434	HEM1_PASMU	P93525 pasteurella
14	31	70.5	448	HEM1_STAAM	Q99ti9 staphylococcus
15	31	70.5	457	LAT_STRCL	Q01767 streptomyces
16	31	70.5	1451	MYM1_HUMAN	P52179 homo sapien
17	31	70.5	1669	CA14_HUMAN	P02462 homo sapien
18	31	70.5	1669	CA14_MOUSE	P02463 mus musculus
19	30	68.2	114	PAND_THEMEA	Q9x037 thermotoga
20	30	68.2	191	TF2D_ACECL	P46272 acetabularia
21	30	68.2	365	YN07_ARCFU	O27977 archaeoglob
22	30	68.2	399	VP47_NPVOP	O10300 orgyia pseu
23	30	68.2	413	HEM1_COXBU	P47846 coxiella bu
24	30	68.2	419	HEM1_VIBCH	Q9kq24 vibrio chol
25	30	68.2	429	PYRC_AERPE	Q9yfi5 aeropyrum p
26	30	68.2	440	V117_FOWPV	Q9j5a5 fowlpox vir
27	30	68.2	443	HEM1_HELMO	P92966 heliobacill
28	30	68.2	455	HEM1_BACSU	P16618 bacillus su
29	30	68.2	505	SYE_CHLPS	Q06560 chlamydia p
30	30	68.2	506	SYE_CHLMU	Q9piu7 chlamydia m
31	30	68.2	506	SYE_CHLTR	O84451 chlamydia t
32	30	68.2	514	PM12_CHLPN	Q923d6 chlamydia p
33	30	68.2	670	UVRB_METAC	Q8tkk3 methanosarc

34 30 68.2 1194 1 DPOL\_VZVD P09252 varicella-z  
35 29 65.9 75 1 LSM6\_SCHPO Q9uul1 schizosacch  
36 29 65.9 113 1 Y586\_DEIRA Q9rws9 deinococcus  
37 29 65.9 186 1 VMA2\_BRSVA P29792 bovine resp  
38 29 65.9 186 1 VMA2\_ORSVW Q84132 ovine respi  
39 29 65.9 194 1 VMA2\_HRSVA P04435 human respi  
40 29 65.9 218 1 TRPF\_SYNY3 P74435 synechocyst  
41 29 65.9 257 1 MVH8\_RAT P04462 rattus norv  
42 29 65.9 270 1 YBJ5\_YEAST P38172 saccharomyc  
43 29 65.9 297 1 YXNG\_BACSU P24810 bacillus su  
44 29 65.9 298 1 VBL1\_CLVK P14981 cassava lat  
45 29 65.9 298 1 VBL1\_CLVN P14971 cassava lat

ALIGNMENTS

RESULT 1  
VE6\_HPV45  
ID VE6\_HPV45 STANDARD; PRT; 158 AA.  
AC P21735;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE E6 protein.  
GN E6.  
OS Human papillomavirus type 45.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10593;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; PubMed=8205838;  
RA Delliou H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kaplan J.B., Burk R.D.;  
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.  
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
CC STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X74479; CAA52573.1; -;  
CC EMBL; M38198; AAA46973.1; -;  
CC PIR; S36561; S36561.  
CC InterPro: IPR001334; E6.  
CC Pfam: PF00518; E6; 1.  
CC Early protein; DNA-binding; Nuclear protein; Zinc-finger;  
CC Oncogene.  
KW  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 105 141 POTENTIAL.  
FT ZN\_FING 10 10 R -> P (IN REF. 2).  
FT CONFLICT 30 30 I -> N (IN REF. 2).  
FT CONFLICT 118 118 R -> A (IN REF. 2).  
SQ SEQUENCE 158 AA; 18897 MW; F1CF10DD33AA4C3E CRC64;  
  
Query Match 100.0%; Score 44; DB 1; Length 158;  
Best Local Similarity 100.0%; Pred. No. 0.013; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;  
  
QY 1 ATLERTEVY 9  
|||||||

```

DB 37 ATLERTEVY 45
RESULT 2
HEM1_DEIRA
ID HEM1_DEIRA STANDARD; PRT; 350 AA.
AC Q9RRES;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTR).
GN HEMA OR DR2547.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1".
RL Science 286:1571-1577(1999).
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90400; BAA31845.1; -
DR PIR; E36779; W6WL58.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
SQ SEQUENCE 149 AA; 17794 MW; 79B3DCC95831B158 CRC64;
Query Match 75.0%; Score 33; DB 1; Length 149;
Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 2 TLTERTEVY 9
Db 36 TLQRSEVY 43
RESULT 4
HEM1_PAEMA
ID HEM1_PAEMA STANDARD; PRT; 437 AA.
AC O69108;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTR).
GN HEMA.
OS Paenibacillus macerans (Bacillus macerans).
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OX NCBI_TaxID=44252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 24;
RA Johansson P., Hederstedt L.;
RT "Organization of genes for tetrapyrrole biosynthesis in Gram-positive
RT bacteria".
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF064061; AAC18585.1; -
Query Match 77.3%; Score 34; DB 1; Length 350;
Best Local Similarity 77.8%; Pred. No. 5.3;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATLERTEVY 9
Db 52 ATCNRTVEY 60
RESULT 3
VE6_HPV58
ID VE6_HPV58 STANDARD; PRT; 149 AA.
AC P26553;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE E1-JUN-1994 (Rel. 29, Last annotation update)
GN E6 protein.
DE E6.

```



DR HSP; Q42843; I29.  
DR InterPro: IPR000343; Glutr.  
DR Pfam: PF00745; Glutr; 1.  
DR TIGRFAMS: TIGR01035; hema; 1.  
DR PROSITE: PS00747; GLUTR; 1.  
KW Porphyrin biosynthesis; Oxidoreductase; NADP.  
FT ACT\_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).  
FT ACT\_SITE 99 99 BASE (BY SIMILARITY).  
SQ SEQUENCE 437 AA; 49394 MW; 51EAD60077424FB2 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 437;  
Best Local Similarity 66.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATLERTEYV 9  
II IIII  
DB 48 ATCNRTIY 56

RESULT 5  
ID V66\_HPV68 STANDARD; PRT; 158 AA.  
AC P54667;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE E6 protein.  
GN E6.  
OS Human papillomavirus type 68.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=45240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97060129; PubMed=8904450;  
RA Longuet M., Beaudenon S., Orth G.;  
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,  
related to the potentially oncogenic HPV39.";  
RL J. Clin. Microbiol. 34:738-744(1996).  
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X67160; CAA47632.1; -;  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 105 141 POTENTIAL.  
SQ SEQUENCE 158 AA; 18796 MW; 46B37939CFBA6596 CRC64;  
Query Match 72.7%; Score 32; DB 1; Length 158;  
Best Local Similarity 85.7%; Pred. No. 6;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEYV 9  
I:IIII  
DB 39 LORTEYV 45

RESULT 6  
ID V66\_HPVME STANDARD; PRT; 158 AA.  
AC P27962;  
DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE E6 protein.  
GN E6  
OS Human papillomavirus type ME180.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91374616; PubMed=1716694;  
RA Reuter S., Delius H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;  
RT "Characterization of a novel human papillomavirus DNA in the cervical  
carcinoma cell line ME180.";  
RL J. Virol. 65:5564-5568(1991).  
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M73258; -; NOT\_ANNOTATED\_CDS.  
DR PIR; C40509; W6WLPK.  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 106 141 POTENTIAL.  
SQ SEQUENCE 158 AA; 18738 MW; 2B1F406B563F05FC CRC64;

Query Match 72.7%; Score 32; DB 1; Length 158;  
Best Local Similarity 85.7%; Pred. No. 6;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEYV 9  
I:IIII  
DB 39 LORTEYV 45

RESULT 7  
ID GMDL\_CAEEL STANDARD; PRT; 368 AA.  
AC Q18801;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable GDP-mannose 4,6 dehydratase 1 (EC 4.2.1.47) (GDP-D-mannose  
dehydratase) (GMD).  
GN C53B4.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Berks M.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: CONVERSION OF GDP-D-MANNOSE TO GDP-4-KEFO-6-D-  
DEOXYMANNOSE (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: GDP-mannose = GDP-4-dehydro-6-deoxy-D-mannose  
+ H(2)O.  
CC -!- COFACTOR: NADP(+) (BY SIMILARITY).  
CC -!- PATHWAY: Conversion of GDP-mannose to GDP-fucose; first step.  
CC -!- SIMILARITY: BELONGS TO THE GDP-MANNOSE 4,6-DEHYDRATASE FAMILY.

DT 01-AUG-1992 (Rel. 23, Created)

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: Z68215; CAA92455.2; -  
 DR HSSP: P32054; IDB3  
 DR WormPep: C53B4.7; CE23608.  
 KW Lyase; NADP.  
 SQ SEQUENCE 368 AA; 41309 MW; 9B6E737154AAE70C CRC64;

Query Match 72.7%; Score 32; DB 1; Length 368;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
 :|||  
 Db 95 STIEPTVY 103

RESULT 8  
 UVRB\_STRPN STANDARD; PRT: 662 AA.  
 AC Q54986;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Excinuclease ABC subunit B.  
 GN UVRB OR UVS402 OR SPI238.  
 OS Streptococcus pneumoniae.  
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
 Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
 McDonald L.A., Feldblum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
 Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 "Complete genome sequence of a virulent isolate of Streptococcus  
 pneumoniae.";  
 RT Science 293:498-506(2001).  
 RL

CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT  
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS  
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES  
 CC THE ATPASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED  
 CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND  
 CC TO UV-IRRADIATED DUPLEX DNA (BY SIMILARITY).  
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRB, UVRB AND UVRB.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: M80215; AAA27020.1; ALT\_INIT.  
 DR EMBL: AE007423; AAK75344.1; -  
 DR HSSP: P56981; ID9X.  
 DR TIGR: SPI238; -

DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR004807; UvrB.  
 DR InterPro: IPR001943; UvrB/C.  
 DR Pfam: PF00271; Helicase\_C; 1.  
 DR Pfam: PF02151; UVR; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR SMART: SM00490; HELIC; 1.  
 DR TIGRFAMs: TIGR00631; uvrB; 1.  
 DR PROSITE: PS0151; UVR; 1.

KW SOS response; Excision nuclease; DNA repair; ATP-binding;  
 KW Complete proteome.  
 FT NP\_BIND 44 51 ATP (POTENTIAL).  
 FT DOMAIN 626 661 UVR.  
 FT CONFLICT 349 349 I -> M (IN REF. 1).  
 FT CONFLICT 364 364 E -> K (IN REF. 1).

FT SEQUENCE 662 AA; 75733 MW; C2EB617ED537BDD1 CRC64;  
 Query Match 72.7%; Score 32; DB 1; Length 662;  
 Best Local Similarity 85.7%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
 :|||  
 Db 486 TLERTEI 492

RESULT 9  
 UVRB\_STRPY STANDARD; PRT: 663 AA.  
 AC Q99ZA5;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Excinuclease ABC subunit B.  
 GN UVRB OR SPI1314.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;  
 RX MEDLINE=21192684; PubMed=11296296;  
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,  
 Primeaux C., Sezate S., Suvoarov A.N., Kenton S., Lai H.S., Lin S.P.,  
 Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
 Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 RL

CC -1- FUNCTION: The ABC excision nuclease is a DNA repair enzyme that  
 CC catalyzes the excision reaction of UV-damaged nucleotide segments  
 CC producing oligomers having the modified base(s). UvrB stimulates  
 CC the ATPase activity of uvrA in the presence of UV-irradiated  
 CC double-stranded DNA. It also enhances the ability of uvrA to bind  
 CC to UV-irradiated duplex DNA (By similarity).  
 CC -1- SUBUNIT: Consists of three subunits; uvrA, uvrB and uvrC.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE006570; AAK34157.1; -.  
HSSP; P56981; 1D9X.  
InterPro; IPR001410; DEAD.  
InterPro; IPR001650; Helicase\_C.  
InterPro; IPR004807; UvrB/C.  
Pfam; PF00271; helicase\_C; 1.  
Pfam; PF02151; UVR; 1.  
SMART; SM00487; DEXDC; 1.  
SMART; SM00490; HELICC; 1.  
TIGRFAMS; TIGR00631; uvrB; 1.  
PROSITE; PS50151; UVR; 1.  
SOS response; Excision nuclease; DNA repair; ATP-binding;  
Complete proteome.  
NP\_BIND 44 51 ATP (POTENTIAL).  
DOMAIN 627 662 UVR.  
SEQUENCE 663 AA; 75654 MW; 999CB73CD049C53A CRC64;

Query Match 72.7%; Score 32; DB 1; Length 663;  
Best Local Similarity 85.7%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
Db 486 TLERTEI 492  
|||||

RESULT 10  
ID UVRB\_LACLA STANDARD; PRT; 692 AA.  
AC Q9C106;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Exonuclease ABC subunit B.  
GN UVRB OR LL0562.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=11337471;  
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,  
RA Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
lactis ssp. lactis IL1403";  
RL Genome Res. 11:731-753(2001).  
CC -!- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT  
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS  
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES  
CC THE ATPASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED  
CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND  
CC TO UV-IRRADIATED DUPLEX DNA (BY SIMILARITY).  
CC -!- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRB, UVRB AND UVRB.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE UVRB FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 UVR DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE006290; AAK04660.1; -.  
HSSP; P56981; 1D9X.

InterPro; IPR001410; DEAD.  
InterPro; IPR001650; Helicase\_C.  
InterPro; IPR004807; UvrB.  
InterPro; IPR001943; UvrB/C.  
Pfam; PF00271; helicase\_C; 1.  
Pfam; PF02151; UVR; 1.  
SMART; SM00487; DEXDC; 1.  
SMART; SM00490; HELICC; 1.  
TIGRFAMS; TIGR00631; uvrB; 1.  
PROSITE; PS50151; UVR; 1.  
SOS response; Excision nuclease; DNA repair; ATP-binding;  
Complete proteome.  
NP\_BIND 45 52 ATP (POTENTIAL).  
DOMAIN 656 691 UVR.  
SEQUENCE 692 AA; 79131 MW; 9585B28875E7C402 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 692;  
Best Local Similarity 85.7%; Pred. No. 32;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEV 8  
Db 487 TLERTEI 493  
|||||

RESULT 11  
ID PTP6\_DROME STANDARD; PRT; 1462 AA.  
AC P16620;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Protein-tyrosine phosphatase DPTP precursor (EC 3.1.3.48) (Protein-  
DE tyrosine-phosphate phosphohydrolase).  
GN PTP69D OR DPTP.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90046860; PubMed=2554325;  
RA Streuli M., Krueger N.X., Tsai A.Y.M., Saito H.;  
RT "A family of receptor-linked protein tyrosine phosphatases in humans  
and Drosophila";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:8698-8702(1989).  
CC -!- FUNCTION: IT IS POSSIBLE THAT DPTP IS A CELL ADHESION RECEPTOR.  
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein  
CC tyrosine + phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M27699; AAA28842.1; -.  
PIR; B36182; B36182.  
HSSP; P18052; 1YFO.  
FlyBase; FBgn0014007; Ptp69D.  
InterPro; IPR003961; FN\_III.  
InterPro; IPR003006; Ig\_MHC.  
InterPro; IPR003598; Ig\_c2.  
InterPro; IPR003600; Ig\_like.  
InterPro; IPR000387; Tyr\_phosphatase.  
InterPro; IPR000242; Tyr\_PP.

DR Pfam; PF00041; fn3; 3.  
 DR Pfam; PF00047; Ig; 2.  
 DR PRINTS; PR00700; PRTYHPHPTASE.  
 DR SMART; SM00060; FN3; 3.  
 DR SMART; SM00410; IG-like; 1.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00194; PTPC; 2.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 2.  
 DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; 2.  
 DR PROSITE; PS00055; TYR\_PHOSPHATASE\_PTP; 2.  
 DR Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane;  
 KW Cell adhesion; Immunoglobulin domain; Repeat.  
 FT SIGNAL 1 23  
 FT CHAIN 24 1462 PROTEIN-TYROSINE PHOSPHATASE DPTP.  
 FT DOMAIN 24 805 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 806 823 POTENTIAL.  
 FT DOMAIN 824 1462 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 38 119 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 147 221 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 332 437 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 438 538 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 912 1165 PROTEIN-TYROSINE PHOSPHATASE 1.  
 FT DOMAIN 1208 1459 PROTEIN-TYROSINE PHOSPHATASE 2.  
 FT ACT\_SITE 1097 1097 BY SIMILARITY.  
 FT ACT\_SITE 1391 1391 BY SIMILARITY.  
 FT DISULFID 45 112 POTENTIAL.  
 FT DISULFID 154 214 POTENTIAL.  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 516 516 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 701 701 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 755 755 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1462 AA; 167411 MW; F8091D69E88230EB CRC64;  
 Query Match 72.7%; Score 32; DB 1; Length 1462;  
 Best Local Similarity 87.5%; Pred. No. 74;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 TLERTEVY 9  
 Db 198 TLEFTEVY 205  
 RESULT 12  
 GMD2\_CAEEL STANDARD; PRT; 382 AA.  
 ID GMD2\_CAEEL  
 AC 045583;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Probable GDP-mannose 4,6 dehydratase 2 (EC 4.2.1.47) (GDP-D-mannose  
 dehydratase) (GMD).  
 GN F56H6.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Kershaw J.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: CONVERSION OF GDP-D-MANNOSE TO GDP-4-KETO-6-D-  
 DEOXYMANNOSE (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: GDP-mannose = GDP-4-dehydro-6-deoxy-D-mannose  
 + H(2)O.  
 CC -!- COFACTOR: NADP(+) (BY SIMILARITY).  
 CC -!- PATHWAY: Conversion of GDP-mannose to GDP-fucose; first step.  
 CC -!- SIMILARITY: BELONGS TO THE GDP-MANNOSE 4,6-DEHYDRATASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z81553; CAB04494.1; -  
 DR HSSP; P32054; 1DB3.  
 DR WormPep; F56H6.5; CE16133.  
 KW Lyase; NADP.  
 SQ SEQUENCE 382 AA; 43071 MW; 1A3437F87872B457 CRC64;  
 Query Match 70.5%; Score 31; DB 1; Length 382;  
 Best Local Similarity 55.6%; Pred. No. 27;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATLE RTEVY 9  
 Db 109 STIEPTEIY 117  
 RESULT 13  
 HEMI\_PASMU STANDARD; PRT; 434 AA.  
 ID HEMI\_PASMU  
 AC P95525;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glutamyl-tRNA reductase (EC 1.2.1.1-) (GluTR).  
 GN HEMA OR GLTX1 OR PM0684.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Pasteurella.  
 OX NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Castrillon R.T.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pm70;  
 RX MEDLINE=21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-  
 semialdehyde + NADP(+) + tRNA(Glu).  
 CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.  
 CC -!- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----



```
CC LYSINE OR L-ORNITHINE TO ALPHA-KETOGLUTARATE. OXALACETATE AND
CC PYRUVATE CAN ALSO BE USED AS ACCEPTORS OF THE AMINO GROUP BUT WITH
CC VERY LOW EFFICIENCY. OPTIMAL ACTIVITY AT PH 7.0-7.5 AND 30 DEGREES
CC CELSIUS.
CC -!- CATALYTIC ACTIVITY: L-lysine + 2-oxoglutarate = 2-aminoadipate 6-
CC semialdehyde + L-glutamate.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: BETA-LACTAM BIOSYNTHESIS: FIRST STEP IN CEPHAMYCIN C
CC SYNTHESIS.
CC -!- SUBUNIT: ACTIVE AS EITHER A MONOMER OR A HOMOPOLYMER.
CC -!- MISCELLANEOUS: PRESENT ONLY IN BETA-LACTAM PRODUCING
CC ACTINOMYCETES.
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M64834; AAA26777.1; -.
CC EMBL; U12015; AAB39899.1; -.
CC PIR; A38169; A38169.
CC InterPro; IPR000954; Aminotran_3.
CC Pfam; PF00202; aminotran_3.1.
CC PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
CC Transferase; Aminotransferase; Antibiotic biosynthesis;
KW Pyridoxal phosphate.
KW BINDING 304 304 PYRIDOXAL PHOSPHATE (POTENTIAL).
FT CONFLICT 123 123 R -> P (IN REF. 2).
FT CONFLICT 452 452 P -> A (IN REF. 2).
SQ SEQUENCE 457 AA; 49974 MW; 94C660C027795FAF CRC64;

Query Match 70.5%; Score 31; DB 1; Length 457;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
Db 351 TIERTQVF 358

Search completed: February 6, 2003, 14:28:02
Job time : 13 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:27:37 ; Search time 29 Seconds  
(without alignments)  
63.946 Million cell updates/sec

Title: US-09-641-528a-31040  
Perfect score: 44  
Sequence: 1 ATLETERVEY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	158	4 Q9Y4Y4	Q9Y4Y4 homo sapien
2	44	100.0	158	12 O10608	O10608 human papil
3	40	90.9	212	16 Q92SQ2	Q92SQ2 rhizobium m
4	35	79.5	212	16 Q987R6	Q987R6 rhizobium l
5	33	75.0	149	12 Q91ZG5	Q91ZG5 human papil
6	33	75.0	149	12 Q91ZG6	Q91ZG6 human papil
7	33	75.0	149	12 Q8QSE7	Q8QSE7 human papil
8	33	75.0	149	12 Q8QHQ3	Q8QHQ3 human papil
9	33	75.0	157	12 Q9WHG0	Q9WHG0 human papil
10	33	75.0	1264	10 Q9XIG5	Q9XIG5 arabidopsis
11	33	72.7	90	16 Q9PAU4	Q9PAU4 xylella fas
12	32	72.7	90	16 Q9PAU0	Q9PAU0 xylella fas
13	32	72.7	149	12 Q90723	Q90723 human papil
14	32	72.7	162	4 Q901Z9	Q901Z9 homo sapien
15	32	72.7	215	16 Q9KMZ6	Q9KMZ6 vibrio chol
16	32	72.7	262	8 Q9B8D8	Q9B8D8 candida alb

17	32	72.7	383	10 Q9SP62	Q9SP62 impatiens b
18	32	72.7	384	5 Q8T8N2	Q8T8N2 caenorhabdi
19	32	72.7	399	5 Q8T8N3	Q8T8N3 caenorhabdi
20	32	72.7	455	2 O66085	O66085 lactococcus
21	32	72.7	911	16 Q8RFJ1	Q8RFJ1 fusobacteri
22	32	72.7	1191	11 P97690	P97690 rattus norv
23	32	72.7	1217	4 O60464	O60464 homo sapien
24	32	72.7	1217	11 Q9QUS3	Q9QUS3 mus musculu
25	32	72.7	1218	6 Q97594	Q97594 bos taurus
26	32	72.7	1428	5 Q9VU03	Q9VU03 drosophila
27	32	72.7	3012	5 Q97205	Q97205 leishmania
28	31	70.5	135	5 Q9VG36	Q9VG36 drosophila
29	31	70.5	167	16 Q8YQD9	Q8YQD9 anabaena sp
30	31	70.5	183	16 P74169	P74169 synechocyst
31	31	70.5	225	6 Q28271	Q28271 canis famil
32	31	70.5	226	11 Q99LQ8	Q99LQ8 mus musculu
33	31	70.5	229	4 Q9NYC5	Q9NYC5 homo sapien
34	31	70.5	245	16 Q9KDB4	Q9KDB4 bacillus ha
35	31	70.5	288	10 Q9LVC8	Q9LVC8 arabidopsis
36	31	70.5	349	13 Q9DEB8	Q9DEB8 gallus gall
37	31	70.5	372	5 Q965S5	Q965S5 caenorhabdi
38	31	70.5	427	9 Q8SD32	Q8SD32 pseudomonas
39	31	70.5	438	12 Q8VSE3	Q8VSE3 ndelle viru
40	31	70.5	501	2 O07068	O07068 burkholderi
41	31	70.5	582	17 Q980Q0	Q980Q0 sulfolobus
42	31	70.5	703	16 Q8XTK8	Q8XTK8 raietonia s
43	31	70.5	717	5 Q961F0	Q961F0 drosophila
44	31	70.5	747	5 Q95Q88	Q95Q88 caenorhabdi
45	31	70.5	844	16 Q92S58	Q92S58 rhizobium m

ALIGNMENTS

RESULT 1

Q9Y4Y4  
ID Q9Y4Y4 PRELIMINARY; PRT; 158 AA.  
AC Q9Y4Y4;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE E6 protein.  
GN HPV45 E6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sastre-Garau X., Favre M., Couturier J., Orth G.;  
RT "Distinct patterns of alteration of myc genes associated with  
RT integration of human papillomavirus type 16 oe type 45 in two genital  
RT tumours."  
RL J. Gen. Virol. 81:198-199(2000).  
DR EMBL; AJ242956; CAB44706.1; .  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
SQ SEQUENCE 158 AA; 18914 MW; A61AEF98390AEEB3 CRC64;

Query Match 100.0%; Score 44; DB 4; Length 158;  
Best Local Similarity 100.0%; Pred. No. 0.089;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLETERVEY 9  
Db 37 ATLETERVEY 45

RESULT 2

O10608  
ID O10608 PRELIMINARY; PRT; 158 AA.  
AC O10608;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)

```

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Oncoprotein E6.
GN E6.
OS Human papillomavirus type 45.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICA HPV45 VARIANT;
RA Saxtre-Garau X., Favre M., Couturier J., Orth G.;
RT "Distinct patterns of alteration of myc genes associated with
RL integration of HPV16 or HPV45 DNA in two genital tumors.";
RL J. Gen. Virol. 0:0-0(0).
DR EMBL: Y13218; CAA73660.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 158 AA; 18914 MW; A61AEF98390AE83 CRC64;

Query Match 100.0%; Score 44; DB 12; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLETEVY 9
Db 37 ATLETEVY 45

RESULT 3
Q92SQ2 PRELIMINARY; PRT; 212 AA.
AC Q92SQ2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative MTA/SAH nucleosidase P46 INCLUDES: 5'-methylthioadenosine
DE nucleosidase and S-adenosylhomocysteine nucleosidase protein
DE (EC 3.2.2.16).
GN PFS OR R00315 OR SMC00395.
OS Rhizobium meliloti (Sinorhizobium meliloti). Rhizobiaceae group;
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ransperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RW EMBL: AL591783; CAC41752.1; -.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 212 AA; 22431 MW; C7AB99BDA2DE6158 CRC64;

Query Match 90.9%; Score 40; DB 16; Length 212;
Best Local Similarity 88.9%; Pred. No. 0.94;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLETEVY 9
Db 73 ATLEQTEVY 81

RESULT 4
Q987R6 PRELIMINARY; PRT; 212 AA.
ID Q987R6
AC Q987R6;

```

```

DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE MTA/SAH nucleosidase.
GN MUR6947.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003010; BAB53134.1; -.
KW Complete proteome.
SQ SEQUENCE 212 AA; 22573 MW; 535012642A0DF7EA CRC64;

Query Match 79.5%; Score 35; DB 16; Length 212;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
Db 74 TLEQTEVY 81

RESULT 5
Q91ZG5 PRELIMINARY; PRT; 149 AA.
AC Q91ZG5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 58.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGHK3;
RA Chan P.K.S., Lam C.W., Li W.H., Chan M.Y.M., Cheung J.L.K.,
RA Cheng A.F.;
RT "Nucleotide sequence variation of the E6 gene of human papillomavirus
RT type 58 from Chinese women with normal cervixes and cancerous
RT lesions.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RW EMBL: AF234531; AAF60307.1; -.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6; 1.
SQ SEQUENCE 149 AA; 17863 MW; 4C6A61AD9CEC302D CRC64;

Query Match 75.0%; Score 33; DB 12; Length 149;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
Db 36 TLQRSEVY 43

RESULT 6
Q91ZG6 PRELIMINARY; PRT; 149 AA.
ID Q91ZG6

```



1

```

Db 37 SVLQTEVY 45
RESULT 10
Q9XIG5 ID Q9XIG5 PRELIMINARY; PRT; 1264 AA.
AC Q9XIG5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE T10P12.3 protein.
DE T10P12.3.
GN T10P12.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Lueros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007203; AAD39272.1; -
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00385; chromo; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; rvt; 1.
DR SMART; SM00298; CHROMO; 1.
DR PROSITE; PS50013; CHROMO_2; 1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1264 AA; 144919 MW; FC4E32CD56B78DBA CRC64;

Query Match 75.0%; Score 33; DB 10; Length 1264;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8
:|||||:
Db 62 STLERTEI 69

RESULT 11
Q9PAU4 ID Q9PAU4 PRELIMINARY; PRT; 90 AA.
AC Q9PAU4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Xf2401.
GN Xf2401.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Paixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE004049; AAF85200.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 9720 MW; 5E2FD2193395B279 CRC64;

Query Match 72.7%; Score 32; DB 16; Length 90;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
|||||
Db 12 ATLDRAEAY 20

RESULT 12
Q9PAU0 ID Q9PAU0 PRELIMINARY; PRT; 90 AA.
AC Q9PAU0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Xf2405.
GN Xf2405.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Paixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,

```

```
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AE004049; AAF85204.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 9863 MW; 52D8E8A5D1A4F75 CRC64;

Query Match 72.7%; Score 32; DB 16; Length 90;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
Db 12 ATLDRAEAY 20

RESULT 13
O90723 PRELIMINARY; PRT; 149 AA.
AC O90723;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF E6.
OS Human papillomavirus type 67.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OC NCBI_TaxID=37120;
RN [1]
RP SEQUENCE FROM N.A.
RA Kirii Y., Iwamoto S., Matsukura T.;
RT "Complete DNA sequence of human Papillomavirus type 67 isolated from a
RT vaginal intraepithelial neoplasia.";
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99073695; PubMed=9857984;
RA Kirii Y., Matsukura T.;
RT "Nucleotide sequence and phylogenetic classification of
RT humanpapillomavirus type 67.";
RL Virus Genes 17:117-121(1998).
DR EMBL; D21208; BAA28852.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
SQ SEQUENCE 149 AA; 17832 MW; A4F28AC1BE75B197 CRC64;

Query Match 72.7%; Score 32; DB 12; Length 149;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
Db 36 TLDREVY 43

RESULT 14
Q9UI29 PRELIMINARY; PRT; 162 AA.
AC Q9UI29;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cellular DNA/human papillomavirus proviral DNA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CERVIX;
RX MEDLINE=91374616; PubMed=1716694;
RA Reuter S., Delius H., Kahn T., Hofmann B., zur Hausen H., Schwarz E.;
```

```
RT "Characterization of a novel human papillomavirus DNA in the cervical
RT carcinoma cell line ME180.";
RL J. Virol. 65:5564-5568(1991).
DR EMBL: M73258; AAF14011.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
SQ SEQUENCE 162 AA; 19167 MW; 2E40C128FEF7902C CRC64;

Query Match 72.7%; Score 32; DB 4; Length 162;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 43 LQTEVY 49

RESULT 15
Q9KMZ6 PRELIMINARY; PRT; 215 AA.
AC Q9KMZ6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein VCA0170.
GN VCA0170.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OC NCBI_TaxID=686;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004357; AAF96083.1; -.
DR TIGR; VCA0170; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 215 AA; 24121 MW; BDA345490673FFC4 CRC64;

Query Match 72.7%; Score 32; DB 16; Length 215;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 144 LERTELY 150

Search completed: February 6, 2003, 14:29:41
Job time : 30 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:28:07 : Search time 139 seconds  
(without alignments)  
41.745 Million cell updates/sec

Title: US-09-641-528A-31040  
Perfect score: 44  
Sequence: 1 ATLETRTEVY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main : \*  
1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	158	1	PCT-US00-33549-51
2	34	77.3	350	27	US-60-360-039-678
3	34	77.3	541	27	US-60-242-679-870
4	34	77.3	611	27	US-60-212-659-524
5	34	77.3	611	27	US-60-230-435-1581
6	33	75.0	437	21	US-09-791-537-35359

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	158	1	PCT-US00-33549-51
2	34	77.3	350	27	US-60-360-039-678
3	34	77.3	541	27	US-60-242-679-870
4	34	77.3	611	27	US-60-212-659-524
5	34	77.3	611	27	US-60-230-435-1581
6	33	75.0	437	21	US-09-791-537-35359

Sequence 51, Appl  
Sequence 678, App  
Sequence 870, App  
Sequence 524, App  
Sequence 1581, Ap  
Sequence 35359, A

H-11, M.  
09/641528  
Seg. ID 31040 w/  
Interf

7	33	75.0	469	1	PCT-US01-14827-8950	Sequence 8950, Ap
8	33	75.0	547	1	PCT-US01-14827-9982	Sequence 9982, Ap
9	33	75.0	547	27	US-60-323-349-464	Sequence 464, App
10	33	75.0	715	27	US-60-360-039-9642	Sequence 9642, App
11	33	75.0	1258	21	US-09-708-427-15317	Sequence 15317, A
12	33	75.0	1264	21	US-09-708-427-15316	Sequence 15316, A
13	33	75.0	1316	19	US-09-573-655A-2403	Sequence 2403, Ap
14	33	75.0	1316	19	US-09-573-655B-2403	Sequence 2403, Ap
15	32	72.7	159	1	PCT-US02-03987-5454	Sequence 5454, Ap
16	32	72.7	159	22	US-09-815-242-5454	Sequence 5454, Ap
17	32	72.7	159	24	US-10-072-851-5454	Sequence 5454, Ap
18	32	72.7	159	27	US-60-242-578-821	Sequence 821, App
19	32	72.7	159	27	US-60-253-625-2165	Sequence 2165, App
20	32	72.7	159	27	US-60-257-931-2997	Sequence 2997, App
21	32	72.7	159	27	US-60-269-308-4018	Sequence 4018, App
22	32	72.7	162	1	PCT-US02-03987-12683	Sequence 12683, A
23	32	72.7	162	22	US-09-815-242-12683	Sequence 12683, A
24	32	72.7	162	24	US-10-072-851-12683	Sequence 12683, A
25	32	72.7	253	19	US-09-540-236-2857	Sequence 2857, Ap
26	32	72.7	253	27	US-60-128-476-3815	Sequence 3815, Ap
27	32	72.7	368	21	US-09-791-537-119427	Sequence 119427, Ap
28	32	72.7	383	17	US-09-375-299A-2	Sequence 2, Appli
29	32	72.7	383	20	US-09-638-937-7	Sequence 7, Appli
30	32	72.7	383	26	US-10-224-446-2	Sequence 2, Appli
31	32	72.7	383	27	US-60-097-186-2	Sequence 2, Appli
32	32	72.7	383	27	US-60-142-756-2	Sequence 2, Appli
33	32	72.7	394	15	US-09-134-000-4332	Sequence 4332, Ap
34	32	72.7	455	21	US-09-791-537-15909	Sequence 15909, A
35	32	72.7	514	23	US-09-902-540-10182	Sequence 10182, A
36	32	72.7	661	15	US-09-107-532-3677	Sequence 3677, Ap
37	32	72.7	661	15	US-09-107-532A-3677	Sequence 3677, Ap
38	32	72.7	662	19	US-09-583-110-4571	Sequence 4571, Ap
39	32	72.7	668	21	US-09-791-537-77352	Sequence 77352, A
40	32	72.7	679	15	US-09-107-433-3402	Sequence 3402, Ap
41	32	72.7	1242	1	PCT-US00-05883-911	Sequence 911, App
42	32	72.7	1242	23	US-09-925-299-911	Sequence 911, App
43	32	72.7	1428	20	US-09-614-150-22884	Sequence 22884, A
44	32	72.7	1428	21	US-09-791-537-74000	Sequence 74000, A
45	32	72.7	1428	27	US-60-173-464-18836	Sequence 18836, A

ALIGNMENTS

RESULT 1  
PCT-US00-33549-51  
; Sequence 51, Application PC/TUS0033549  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Human  
; TITLE OF INVENTION: Papillomavirus Using Peptide and Nucleic Acid  
; TITLE OF INVENTION: Compositions  
; FILE REFERENCE: 018623-016110PC  
; CURRENT APPLICATION NUMBER: PCT/US00/33549  
; CURRENT FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR APPLICATION NUMBER: US 09/641,528  
; PRIOR FILING DATE: 2000-08-15  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 51  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Human papillomavirus type 45  
PCT-US00-33549-51

```
Query Match      100.0%; Score 44; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 ATLERTEVY 9
    |||||
Db  37 ATLERTEVY 45

RESULT 2
US-60-360-039-678
; Sequence 678, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 678
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Deinococcus radiodurans
- US-60-360-039-678

Query Match      77.3%; Score 34; DB 27; Length 350;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 ATLERTEVY 9
    |||||
Db  52 ATCNRTEVY 60

RESULT 3
US-60-242-679-870
; Sequence 870, Application US/60242679
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steven Istvan
; APPLICANT: Spier, Eugene
; APPLICANT: Greenberg, Simon
; APPLICANT: Brandenberger, Ralph
; APPLICANT: Wang, Yu
; APPLICANT: Dubman, Alex
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000898-PROV
; CURRENT APPLICATION NUMBER: US/60/242,679
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 2265
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 870
; LENGTH: 541
; TYPE: PRT
; ORGANISM: HUMAN
US-60-242-679-870

Query Match      77.3%; Score 34; DB 27; Length 541;
Best Local Similarity 77.8%; Pred. No. 4.4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 ATLERTEVY 9
    |||||
Db  396 APLERTKVY 404

RESULT 4
```

```
US-60-212-659-524
; Sequence 524, Application US/60212659
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000674
; CURRENT APPLICATION NUMBER: US/60/212,659
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 879
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 524
; LENGTH: 611
; TYPE: PRT
; ORGANISM: HUMAN
US-60-212-659-524

Query Match      77.3%; Score 34; DB 27; Length 611;
Best Local Similarity 77.8%; Pred. No. 5.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 ATLERTEVY 9
    |||||
Db  447 APLERTKVY 455

RESULT 5
US-60-230-435-1581
; Sequence 1581, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1581
; LENGTH: 611
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1581

Query Match      77.3%; Score 34; DB 27; Length 611;
Best Local Similarity 77.8%; Pred. No. 5.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 ATLERTEVY 9
    |||||
Db  447 APLERTKVY 455

RESULT 6
US-09-791-537-35359
; Sequence 35359, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35359
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Paenibacillus macerans
```

US-09-791-537-35359

Query Match 75.0%; Score 33; DB 21; Length 437;  
Best Local Similarity 66.7%; Pred. No. 5.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
| | | | |  
Db 48 ATCNRTLY 56

RESULT 7

PCT-US01-14827-8950  
; Sequence 8950, Application PC/TUS0114827  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-104  
; CURRENT APPLICATION NUMBER: PCT/US01/14827  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/577,408  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 16102  
; SOFTWARE: Custom  
; SEQ ID NO 8950  
; LENGTH: 469  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (427)...(439)  
; OTHER INFORMATION: Pentapeptide repeat proteins (8 copies) domain identified by  
; OTHER INFORMATION: eMATRIX, accession number PF00805D, p-value=3.520e-09, raw score  
; NAME/KEY: DOMAIN  
; LOCATION: (169)...(269)  
; OTHER INFORMATION: K+ channel tetramerisation domain identified by Pfam,  
; OTHER INFORMATION: accession name K\_tetra, E-value=3.8e-22, Pfam score of 87.0  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(469)  
; OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
PCT-US01-14827-8950

Query Match 75.0%; Score 33; DB 1; Length 469;  
Best Local Similarity 66.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
| | | | |  
Db 338 ANLRTDLY 346

RESULT 8

PCT-US01-14827-9982  
; Sequence 9982, Application PC/TUS0114827  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-104  
; CURRENT APPLICATION NUMBER: PCT/US01/14827  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/577,408  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 16102  
; SOFTWARE: Custom  
; SEQ ID NO 9982  
; LENGTH: 547  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (424)...(464)  
; OTHER INFORMATION: Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatur

; OTHER INFORMATION: domain identified by eMATRIX, accession number BL00170B, p-val  
; OTHER INFORMATION: 3.455e-27, raw score of 20.97  
; NAME/KEY: DOMAIN  
; LOCATION: (383)...(475)  
; OTHER INFORMATION: Cyclophilin type peptidyl-prolyl cis-tr domain identified by  
; OTHER INFORMATION: Pfam, accession name pro\_isomerase, E-value=4.2e-38, Pfam scor  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(547)  
; OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
PCT-US01-14827-9982

Query Match 75.0%; Score 33; DB 1; Length 547;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEV 8  
| | | | |  
Db 48 TLERTEV 54

RESULT 9

US-60-323-349-464  
; Sequence 464, Application US/60323349  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Goodrich, Ryle W  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ghosh, Malabika  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Weng, Gezhi  
; APPLICANT: Zhou, Ping  
; APPLICANT: Drmanac, Radoje T  
; TITLE OF INVENTION: Novel Nucleic Acids and Secreted Polypeptides  
; FILE REFERENCE: 808  
; CURRENT APPLICATION NUMBER: US/60/323,349  
; CURRENT FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: US 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: PCT/US00/35017  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: PCT/US01/02623  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: US 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: US 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: PCT/US01/03800  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: US 09/515,126  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: US 09/577,409  
; PRIOR FILING DATE: 2000-05-18  
; PRIOR APPLICATION NUMBER: PCT/US01/04927  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: US 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: PCT/US01/04941  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: US 09/649,167

;; PRIOR FILING DATE: 2000-08-23  
;; PRIOR APPLICATION NUMBER: PCT/US01/08631  
;; PRIOR FILING DATE: 2001-03-30  
;; PRIOR APPLICATION NUMBER: US 09/552,929  
;; PRIOR FILING DATE: 2000-04-18  
;; PRIOR APPLICATION NUMBER: US 09/770,160  
;; PRIOR FILING DATE: 2001-01-26  
;; PRIOR APPLICATION NUMBER: PCT/US01/08656  
;; PRIOR FILING DATE: 2001-04-18  
;; PRIOR APPLICATION NUMBER: US 09/577,408  
;; PRIOR FILING DATE: 2000-05-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/14827  
;; PRIOR FILING DATE: 2001-05-16  
;; NUMBER OF SEQ ID NOS: 478  
;; SOFTWARE: PL\_FL\_genes Version 5.0  
;; SEQ ID NO 464  
;; LENGTH: 547  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)..(547)  
;; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set for  
US-60-323-349-464

Query Match 75.0%; Score 33; DB 27; Length 547;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTEV 8  
Db 48 TLERTEV 54

RESULT 10  
US-60-360-039-9642  
;; Sequence 9642, Application US/60360039  
;; GENERAL INFORMATION:  
;; APPLICANT: Cao, Yongwei  
;; APPLICANT: Chen, Xianfeng  
;; APPLICANT: Goldman, Barry S.  
;; APPLICANT: Hinkie, Gregory J.  
;; APPLICANT: Slater, Steven C.  
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
;; FILE REFERENCE: 38-10(52052)A  
;; CURRENT APPLICATION NUMBER: US/60/360,039  
;; CURRENT FILING DATE: 2002-02-21  
;; NUMBER OF SEQ ID NOS: 47374  
;; SEQ ID NO 9642  
;; LENGTH: 715  
;; TYPE: PRT  
;; ORGANISM: Desulfitobacterium hafniese  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (1)..(715)  
;; OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-9642

Query Match 75.0%; Score 33; DB 27; Length 715;  
Best Local Similarity 66.7%; Pred. No. 1e+03;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
Db 541 ATQEQTEIY 549

RESULT 11  
US-09-708-427-15317  
;; Sequence 15317, Application US/09708427  
;; GENERAL INFORMATION:  
;; APPLICANT: N. ALEXANDROV et al.

;; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
;; TITLE OF INVENTION: THEREBY  
;; FILE REFERENCE: 2750-1243P  
;; CURRENT APPLICATION NUMBER: US/09/708,427  
;; CURRENT FILING DATE: 2000-11-09  
;; NUMBER OF SEQ ID NOS: 85364  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 15317  
;; LENGTH: 1258  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: 1..1258  
;; OTHER INFORMATION: Xaa is any amino acid  
;; NAME/KEY: misc\_feature  
;; LOCATION: 1..1258  
;; OTHER INFORMATION: Ceres Seq. ID 1829032  
US-09-708-427-15317

Query Match 75.0%; Score 33; DB 21; Length 1258;  
Best Local Similarity 75.0%; Pred. No. 2.1e+03;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8  
Db 56 STLTERTEI 63

RESULT 12  
US-09-708-427-15316  
;; Sequence 15316, Application US/09708427  
;; GENERAL INFORMATION:  
;; APPLICANT: N. ALEXANDROV et al.  
;; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
;; TITLE OF INVENTION: THEREBY  
;; FILE REFERENCE: 2750-1243P  
;; CURRENT APPLICATION NUMBER: US/09/708,427  
;; CURRENT FILING DATE: 2000-11-09  
;; NUMBER OF SEQ ID NOS: 85364  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 15316  
;; LENGTH: 1264  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: 1..1264  
;; OTHER INFORMATION: Xaa is any amino acid  
;; NAME/KEY: misc\_feature  
;; LOCATION: 1..1264  
;; OTHER INFORMATION: Ceres Seq. ID 1829031  
US-09-708-427-15316

Query Match 75.0%; Score 33; DB 21; Length 1264;  
Best Local Similarity 75.0%; Pred. No. 2.1e+03;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8  
Db 62 STLTERTEI 69

RESULT 13  
US-09-573-655A-2403  
;; Sequence 2403, Application US/09573655A  
;; GENERAL INFORMATION:  
;; APPLICANT: SOLOVYEV, Victor and TROUKHAN, Maxim  
;; TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptide  
;; FILE REFERENCE: 2750-0876P  
;; CURRENT APPLICATION NUMBER: US/09/573,655A  
;; CURRENT FILING DATE: 2000-05-18



Search completed: February 6, 2003, 14:32:26  
Job time : 142 secs

; NUMBER OF SEQ ID NOS: 3280  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2403  
; LENGTH: 1316  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-573-655A-2403

Query Match 75.0%; Score 33; DB 19; Length 1316;  
Best Local Similarity 75.0%; Pred. No. 2.2e+03;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEV 8  
Db 62 STLERTEI 69

RESULT 14  
US-09-573-655B-2403  
; Sequence 2403, Application US/09573655B  
; GENERAL INFORMATION:  
; APPLICANT: SOLOVIEV, Victor and TROUKHAN, Maxim  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-0876P  
; CURRENT APPLICATION NUMBER: US/09/573,655B  
; CURRENT FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 3281  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2403  
; LENGTH: 1316  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-573-655B-2403

Query Match 75.0%; Score 33; DB 19; Length 1316;  
Best Local Similarity 75.0%; Pred. No. 2.2e+03;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEV 8  
Db 62 STLERTEI 69

RESULT 15  
PCT-US02-03987-5454  
; Sequence 5454, Application PC/TUS0203987  
; GENERAL INFORMATION:  
; APPLICANT: Elittra Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C  
; FILE REFERENCE: ELITRA.028VPC  
; CURRENT APPLICATION NUMBER: PCT/US02/03987  
; CURRENT FILING DATE: 2002-02-02  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 15811  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 5454  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
PCT-US02-03987-5454

Query Match 72.7%; Score 32; DB 1; Length 159;  
Best Local Similarity 75.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEV 8  
Db 95 ATMERTEV 102

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 14:21:41 ; Search time 18 Seconds  
(without alignments)  
40.539 Million cell updates/sec

Title: US-09-641-528A-31040  
Perfect score: 44  
Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 405691 seqs, 81078759 residues

Total number of hits satisfying chosen parameters: 405691

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US05\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	44	100.0	9	5 US-09-641-528-1454	Sequence 1454, Ap
2	44	100.0	9	5 US-09-641-528-15691	Sequence 15691, A
3	44	100.0	9	5 US-09-641-528-23424	Sequence 23424, A
4	44	100.0	9	5 US-09-641-528-31040	Sequence 31040, A
5	44	100.0	9	5 US-09-641-528-37859	Sequence 37859, A
6	44	100.0	9	5 US-09-641-528-44004	Sequence 44004, A
7	44	100.0	9	5 US-09-641-528A-1454	Sequence 1454, Ap
8	44	100.0	9	5 US-09-641-528A-15691	Sequence 15691, A
9	44	100.0	9	5 US-09-641-528A-23424	Sequence 23424, A
10	44	100.0	9	5 US-09-641-528A-31040	Sequence 31040, A
11	44	100.0	9	5 US-09-641-528A-37859	Sequence 37859, A
12	44	100.0	9	5 US-09-641-528A-44004	Sequence 44004, A
13	44	100.0	10	5 US-09-641-528-23439	Sequence 23439, A
14	44	100.0	10	5 US-09-641-528-37904	Sequence 37904, A
15	44	100.0	10	5 US-09-641-528-44038	Sequence 44038, A
16	44	100.0	10	5 US-09-641-528A-23439	Sequence 23439, A
17	44	100.0	10	5 US-09-641-528A-37904	Sequence 37904, A
18	44	100.0	10	5 US-09-641-528A-44038	Sequence 44038, A
19	44	100.0	11	5 US-09-641-528-1455	Sequence 1455, Ap
20	44	100.0	11	5 US-09-641-528-15692	Sequence 15692, A
21	44	100.0	11	5 US-09-641-528-19694	Sequence 19694, A
22	44	100.0	11	5 US-09-641-528-23425	Sequence 23425, A
23	44	100.0	11	5 US-09-641-528-31041	Sequence 31041, A
24	44	100.0	11	5 US-09-641-528-37860	Sequence 37860, A
25	44	100.0	11	5 US-09-641-528A-1455	Sequence 1455, Ap
26	44	100.0	11	5 US-09-641-528A-15692	Sequence 15692, A

ALIGNMENTS

RESULT 1  
US-09-641-528-1454  
; Sequence 1454, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172.705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1454  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-1454  
Query Match 100.0% Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
Db 1 ATLERTEVY 9  
RESULT 2  
US-09-641-528-15691  
; Sequence 15691, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641.528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 15691  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-15691

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
| | | | | | | | |  
Db 1 ATLERTEVY 9

RESULT 3  
US-09-641-528-23424  
; Sequence 23424, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 23424  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-23424

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
| | | | | | | | |  
Db 1 ATLERTEVY 9

RESULT 4  
US-09-641-528-31040  
; Sequence 31040, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528

; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 31040  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-31040

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
| | | | | | | | |  
Db 1 ATLERTEVY 9

RESULT 5  
US-09-641-528-37859  
; Sequence 37859, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 37859  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-37859

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
| | | | | | | | |  
Db 1 ATLERTEVY 9

RESULT 6  
US-09-641-528-44004  
; Sequence 44004, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528  
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44004  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-44004

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
|||||

Db 1 ATLERTEVY 9

## RESULT 7

US-09-641-528A-1454  
; Sequence 1454, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Cellis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1454  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-1454

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
|||||

Db 1 ATLERTEVY 9

## RESULT 8

US-09-641-528A-15691  
; Sequence 15691, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Cellis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; PRIOR APPLICATION NUMBER: US 60/172,705

; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15691  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-15691

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
|||||

Db 1 ATLERTEVY 9

## RESULT 9

US-09-641-528A-23424  
; Sequence 23424, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Cellis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23424  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-23424

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATLERTEVY 9  
|||||

Db 1 ATLERTEVY 9

## RESULT 10

US-09-641-528A-31040  
; Sequence 31040, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Cellis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31040  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-31040

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
Db 1 ATLERTEVY 9

## RESULT 11

US-09-641-528A-37859  
; Sequence 37859, Application US/09641528A  
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 37859  
; LENGTH: 9  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-37859

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
Db 1 ATLERTEVY 9

## RESULT 12

US-09-641-528A-44004

; Sequence 44004, Application US/09641528A  
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44004  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-44004

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
Db 1 ATLERTEVY 9

## RESULT 13

US-09-641-528-23439

; Sequence 23439, Application US/09641528  
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 23439  
; LENGTH: 10  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-23439

Query Match 100.0%; Score 44; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9  
Db 2 ATLERTEVY 10

## RESULT 14

US-09-641-528-37904

; Sequence 37904, Application US/09641528  
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 37904
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-37904
```

```
Query Match          100.0%; Score 44; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATLERTEVY 9
   |||||
Db 2 ATLERTEVY 10
```

```
RESULT 15
US-09-641-528-44038
; Sequence 44038, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2050.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44038
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-44038
```

```
Query Match          100.0%; Score 44; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATLERTEVY 9
   |||||
Db 2 ATLERTEVY 10
```

```
Search completed: February 6, 2003, 14:27:44
Job time : 19 secs
```

**THIS PAGE BLANK (USPTO)**